

STIC-Biotech/ChemLib

106111

From: Salimi, Ali  
Sent: Thursday, October 16, 2003 3:46 PM  
To: STIC-Biotech/ChemLib  
Subject: 08/913,644

Can you please conduct a search including interference search for the application serial no. 08/913,644, for the following:

SEQ ID NO: 2, and 4

Thanks!

Ali Salimi  
AU 1648  
Room 9D07  
Mailbox 8E12

RECEIVED  
OCT 16 2003  
(STIC)

Edward Hart  
Technical Info. Specialist  
STIC/Biotech  
CMI 6B02 Tel: 305-9203

Searcher: \_\_\_\_\_  
Phone: \_\_\_\_\_  
Location: \_\_\_\_\_  
Date Picked Up: 10/13/03  
Date Completed: 10/17/03  
Searcher Prep/Review: \_\_\_\_\_  
Clerical: \_\_\_\_\_  
Online time: \_\_\_\_\_

TYPE OF SEARCH:  
NA Sequences: \_\_\_\_\_  
AA Sequences: \_\_\_\_\_  
Structures: \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Full text: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

VENDOR/COST (where applic.)  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
Questel/Orbit: \_\_\_\_\_  
DRLink: \_\_\_\_\_  
Lexis/Nexis: \_\_\_\_\_  
Sequence Sys.: \_\_\_\_\_  
WWW/Internet: \_\_\_\_\_  
Other (specify): \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**



# STIC SEARCH RESULTS

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*:

Mary Hale, Information Branch Supervisor  
308-4258, CM1-1E01

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 - Circ. Desk



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2003, 10:47:19 ; Search time 62.8512 Seconds

(without alignments)  
1280.394 Million cell updates/sec

Title: US-08-913-644-2

Perfect score: 2733  
Sequence: 1 MLKMRSDMTVTLPPEVSAR.....APSATTSSRPANRVRANK 507

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: A\_Geneseq\_19jun03:\*  
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:\*  
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2733	100.0	507	17	AAW05843
2	2705	99.0	568	21	AAAB23927
3	2705	99.0	568	22	AAAB98429
4	2694	98.6	500	23	AAAO18098
5	2457.5	89.9	536	22	AAAB98442
6	1851	67.7	534	22	AAAB98454
7	1834.5	67.1	504	22	AAAB98436
8	1831	67.0	505	24	AAAG75800
9	1826	66.8	505	20	AAAT08019

10	1826	66.8	505	21	AAV57720	Human papillomavir
11	1826	66.8	505	23	ABB77478	HPV16-L1 1. Human
12	1826	66.8	505	23	ABB77479	HPV16-L1 2. Human
13	1826	66.8	531	22	AAAB98422	Human papillomavir
14	1825	66.8	505	24	AAAB98422	Human papillomavir
15	1824	66.7	505	14	AAAB98407	Human papillomavir
16	1824	66.7	505	19	AAAB98407	HPV16-L1. Bovine p
17	1824	66.7	505	19	AAAB98407	HPV16-L1 gene prot
18	1824	66.7	505	19	AAAB98407	Human papillomavir
19	1824	66.7	505	19	AAAB98407	HPV16-L1 capsid pr
20	1824	66.7	505	20	AAAB98407	Human papillomavir
21	1824	66.7	505	21	AAAB98407	Human papillomavir
22	1818	66.5	505	19	AAAB98407	HPV16-L1 protein -
23	1808	66.2	505	14	AAAB98407	Human papillomavir
24	1808	66.2	505	14	AAAB98407	HPV16-L1. Human pa
25	1808	66.2	505	19	AAAB98407	HPV16-L1 gene prot
26	1808	66.2	505	19	AAAB98407	Human papillomavir
27	1808	66.2	505	20	AAAB98407	Human papillomavir
28	1808	66.2	505	21	AAAB98407	Human papillomavir
29	1801	65.9	531	21	AAAB98407	HPV16-L1 protein.
30	1791	65.5	499	22	AAAB98407	Human papillomavir
31	1789.5	65.5	499	22	AAAB98407	Chimeric biotin-bi
32	1789.5	65.5	499	21	AAAB98407	Chimeric biotin-bi
33	1786	65.3	532	23	AAAB98407	HPV16-L1/57 fusion
34	1785.5	65.3	533	23	AAAB98407	HPV16-L1/57 fusion
35	1783	65.2	533	23	AAAB98407	HPV16-L1/57 fusion
36	1778	65.1	501	21	AAAB98407	HPV16-L1/57 fusion
37	1774	64.9	501	18	AAAB98407	Human HPV6-L1 prot
38	1770	64.7	501	18	AAAB98407	Human HPV6-L1 prot
39	1769	64.7	501	18	AAAB98407	Human HPV6-L1 prot
40	1768.5	64.7	500	18	AAAB98407	Human HPV6-L1 prot
41	1768.5	64.7	500	21	AAAB98407	Human HPV6-L1 prot
42	1768.5	64.7	500	21	AAAB98407	Human HPV6-L1 prot
43	1768.5	64.7	500	22	AAAB98407	Human HPV6-L1 prot
44	1768.5	64.7	500	22	AAAB98407	Human HPV6-L1 prot
45	1768.5	64.7	500	24	AAAB98407	Human papillomavir

#### ALIGNMENTS

RESULT 1  
ID AAW05843 standard; Protein; 507 AA.  
XX  
AC AAW05843:  
XX  
DT 28-JAN-1997 (first entry)  
XX  
DE Human papillomavirus type 18 L1 capsid protein.  
XX  
XX HPV-18: L1 gene; capsid protein; vaccine; diagnosis; vector;  
KW antibody; serotyping; cervix carcinoma.  
XX  
OS Human papillomavirus type 18.  
XX  
PN MO9629413-A2.  
XX  
PD 26-SEP-1996.  
XX  
PE 18-MAR-1996; 96MO-0503649.  
XX  
PR 22-MAR-1995; 95US-0409122.  
XX  
PR 22-MAR-1995; 95US-0408669.  
XX  
PA (MERIT) MERCK & CO INC.  
XX  
PI George HA, Hofmann KU, Jansen KU, Joyce JG, Nepper MP;  
XX WPI: 1996-443188/44.  
XX N-PSDB; AAT40119.  
XX  
PT DNA encoding human papilloma virus 18, esp. L1 and L2 capsid

proteins - and related vectors and antibodies, useful in protective vaccines, for serotyping HPV infections and as therapeutic agents

Claim 2; Fig 1; 46pp; English.

The L1 major capsid protein (AAW05843) of human papillomavirus type 18 (HPV18) has a mol. wt. of 55-60 kDa and is thought to incorporate most of the L2 minor capsid protein (AAW05844). Its amino acid sequence was deduced from a genomic clone (AAT40119) isolated from human cervical carcinoma SW756 cells. Recombinant L1 and L2 proteins can be produced in transformed host cells and used in vaccines for protection against HPV18; this HPV type is associated with invasive carcinomas of the cervix, vagina, vulva and anal canal. Virus-like particles composed of recombinant L1 or L1+L2 can be produced in yeast for use as vaccines.

Sequence 507 AA;

Query Match 100.0%; Score 2733; DB 17; Length 507;  
Best Local Similarity 100.0%; Pred. No. 7e-250;  
Matches 507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MALMRPSDNTVYLPPPSVAVRVNTDDYVTRTSIFYHAGSSRLTVGNPFYFVPAAGGNKQ 60
DB 1 MALMRPSDNTVYLPPPSVAVRVNTDDYVTRTSIFYHAGSSRLTVGNPFYFVPAAGGNKQ 60
QY 61 DIPVSAVQYRVFVQVLPDPKFKGLPDNSIYNPETQRLWMACAGVEIGRQPLGVLSGH 120
DB 61 DIPVSAVQYRVFVQVLPDPKFKGLPDNSIYNPETQRLWMACAGVEIGRQPLGVLSGH 120
QY 121 PFYKKLDDESSHAATSVNSEDVADNVSVYDKQIQLCTLGCAPALGHEHMAKGTACKSRPL 180
DB 121 PFYKKLDDESSHAATSVNSEDVADNVSVYDKQIQLCTLGCAPALGHEHMAKGTACKSRPL 180
QY 121 PFYKKLDDESSHAATSVNSEDVADNVSVYDKQIQLCTLGCAPALGHEHMAKGTACKSRPL 180
DB 121 PFYKKLDDESSHAATSVNSEDVADNVSVYDKQIQLCTLGCAPALGHEHMAKGTACKSRPL 180
QY 181 SQGDCCPLELKNVLEEDGMDVDTGYGAMDFSTLQDKCEVPLDLCQSIKXPDYLOMSAD 240
DB 181 SQGDCCPLELKNVLEEDGMDVDTGYGAMDFSTLQDKCEVPLDLCQSIKXPDYLOMSAD 240
QY 241 PYGDSMFFCLREQLFAHFHFNRACTMGDTVPQSLYIKGTGMRASPGSCVYSPSPSGSIV 300
DB 241 PYGDSMFFCLREQLFAHFHFNRACTMGDTVPQSLYIKGTGMRASPGSCVYSPSPSGSIV 300
QY 301 TSQSOLFENKPYWLKKAAGHNGICMHNQLFTVVDTTRSTNLTICASTQSPVPGQYDATK 360
DB 301 TSQSOLFENKPYWLKKAAGHNGICMHNQLFTVVDTTRSTNLTICASTQSPVPGQYDATK 360
QY 361 FKQYSRHVEEYDLOFIQLCTITLTADVMSYIHSNSSLIEDMNGVPPPTTSLVDYTR 420
DB 361 FKQYSRHVEEYDLOFIQLCTITLTADVMSYIHSNSSLIEDMNGVPPPTTSLVDYTR 420
QY 421 FVQSVAILTCQKDAAPAEKDPYDKLKFVNVDLKEKFSLDLDQYPLGKRFVQAGLRKKPT 480
DB 421 FVQSVAILTCQKDAAPAEKDPYDKLKFVNVDLKEKFSLDLDQYPLGKRFVQAGLRKKPT 480
QY 481 IGPKRKSPASATSSKPAKRVVRARK 507
DB 481 IGPKRKSPASATSSKPAKRVVRARK 507

```

RESULT 2  
AAB23927  
ID AAB23927 standard; protein; 568 AA.

XX AAB23927;

XX 18-JAN-2001 (first entry)

XX Human papillomavirus 18 L1 protein SEQ ID NO:3.

XX Human papillomavirus; HPV16; HPV18; L1 protein; L2 protein; antiviral;  
XX vaccine; immunisation; immune response; infection; diagnosis.

OS Human papillomavirus.

XX WO200054730-A2.  
XX 21-SEP-2000.  
XX 08-MAR-2000; 2000WO-US06017.  
XX 18-MAR-1999; 99US-0125208.  
XX 12-AUG-1999; 99US-0148544.  
XX (HARD ) HARVARD COLLEGE.  
XX (HARR) HARRISON S.  
XX (CHEN/) CHEN X.  
XX Harrison S, Chen X;  
XX WPI; 2000-628165/60.  
XX  
XX Composition comprising multimer of human papillomavirus L1 protein  
XX useful as vaccines against human papillomavirus infection, as  
XX diagnostic tools for detecting presence of the virus in biological  
XX samples -  
XX  
XX Disclosure: Page 30; 31pp; English.

The present invention describes a composition (I) comprising a multimer of human papillomavirus (HPV) L1 protein (II) which is truncated at its amino terminus. (I) has antiviral activity and can be used in the production of a vaccine. (I) comprising a multimer of HPV L1 protein (a T=1 icosahedral particle comprising a pentamer of L1 protein, one of which comprises HPV L2 protein) when administered induces an immune response against the L1 protein in the humans and is thus useful for immunising the humans against HPV. (I) is thus useful as vaccines against human papillomavirus infection, as diagnostic tools for detecting the presence of HPV in biological samples and as tools for mapping receptor interactions. The present sequence represents an HPV18 L1 protein sequence which is used in the exemplification of the present invention.

Sequence 568 AA;

Query Match 99.0%; Score 2705; DB 21; Length 568;  
Best Local Similarity 99.0%; Pred. No. 3.8e-247;  
Matches 502; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

QY 1 MALMRPSDNTVYLPPPSVAVRVNTDDYVTRTSIFYHAGSSRLTVGNPFYFVPAAGGNKQ 60
DB 62 MALMRPSDNTVYLPPPSVAVRVNTDDYVTRTSIFYHAGSSRLTVGNPFYFVPAAGGNKQ 121
QY 61 DIPVSAVQYRVFVQVLPDPKFKGLPDNSIYNPETQRLWMACAGVEIGRQPLGVLSGH 120
DB 122 DIPVSAVQYRVFVQVLPDPKFKGLPDNSIYNPETQRLWMACAGVEIGRQPLGVLSGH 181
QY 121 PFYKKLDDESSHAATSVNSEDVADNVSVYDKQIQLCTLGCAPALGHEHMAKGTACKSRPL 180
DB 182 PFYKKLDDESSHAATSVNSEDVADNVSVYDKQIQLCTLGCAPALGHEHMAKGTACKSRPL 241
QY 181 SQGDCCPLELKNVLEEDGMDVDTGYGAMDFSTLQDKCEVPLDLCQSIKXPDYLOMSAD 240
DB 242 SQGDCCPLELKNVLEEDGMDVDTGYGAMDFSTLQDKCEVPLDLCQSIKXPDYLOMSAD 301
QY 241 PYGDSMFFCLREQLFAHFHFNRACTMGDTVPQSLYIKGTGMRASPGSCVYSPSPSGSIV 300
DB 302 PYGDSMFFCLREQLFAHFHFNRACTMGDTVPQSLYIKGTGMRASPGSCVYSPSPSGSIV 361
QY 301 TSQSOLFENKPYWLKKAAGHNGICMHNQLFTVVDTTRSTNLTICASTQSPVPGQYDATK 360
DB 362 TSQSOLFENKPYWLKKAAGHNGICMHNQLFTVVDTTRSTNLTICASTQSPVPGQYDATK 421
QY 361 FKQYSRHVEEYDLOFIQLCTITLTADVMSYIHSNSSLIEDMNGVPPPTTSLVDYTR 420
DB 422 FKQYSRHVEEYDLOFIQLCTITLTADVMSYIHSNSSLIEDMNGVPPPTTSLVDYTR 481
QY 421 FVQSVAILTCQKDAAPAEKDPYDKLKFVNVDLKEKFSLDLDQYPLGKRFVQAGLRKKPT 480

```

```

DB      482 FVQSVAITCQKDAAPAEKNDPYDKLKFVWVDLKEKFSLDLDQYPLGRKFLVQAGLRKPT 541
QY      481 IGPKRKRSAPSATTTSSSKPAKRVRVRARK 507
DB      542 IGPKRKRSAPSATTTSSSKPAKRVRVRARK 568

RESULT 3
AAB98429
ID      AAB98429 standard; Protein; 568 AA.
AC      AAB98429;
DE      22-AUG-2001 (first entry)
DE      Human papillomavirus protein HPV18 L1.
KW      Human papillomavirus; human leukocyte antigen; HLA; immune response;
KW      HPV; epitope; T cell; identification; vaccine; infection; genital wart;
KW      neoplastic growth; antiviral.
OS      Human papillomavirus.
PN      WO200141799-A1.
PD      14-JUN-2001.
PF      11-DEC-2000; 2000MO-US33549.
PR      10-DEC-1999; 990S-0172705.
PR      15-AUG-2000; 20000S-0641528.
PA      (EPTM-) EPTMONE INC.
PI      Setd A, Sidney J, Southwood S, Chesnut R, Cells E, Grey HM;
DR      WPI: 2001-381497/40.
PT      An isolated human papilloma virus (HPV) epitope, useful in vaccines for
PT      treating HPV infections -
XX      Disclosure: Page 22; 756pp; English.
PS      The present invention describes an isolated prepared human papillomavirus
CC      (HPV) epitope (I). (I) has antiviral activity, and can be used in
CC      vaccine production. Peptides and corresponding nucleic acid compositions.
CC      from the present invention are useful for stimulating an immune response
CC      to HPV by stimulating the production of CTL or HTL responses,
CC      specifically in the treatment or prophylaxis of HPV infection, in persons
CC      who have not manifested symptoms e.g. genital warts or neoplastic growth.
CC      The peptides can also be used in a tetramer staining assay to assess
CC      peripheral blood mononuclear cells for the presence of antigen-specific
CC      CTLs following exposure to a pathogen or immunogen, and as reagents to
CC      evaluate immune recall responses or evaluate the efficacy of a vaccine.
CC      The vaccine compositions are useful for removing warts or treating HPV
CC      infections. The epitopes for inclusion in viral or tumour-associated
CC      be selected from conserved regions of viral or epitope-base vaccine may
CC      antigens, which reduces the likelihood of escape mutants, also
CC      immunosuppressive epitopes that may be present in whole antigens can be
CC      avoided with the use of epitope-base vaccines. An additional advantage
CC      is the ability to combine selected epitopes (CTL and HTL) and to modify
CC      the composition of the epitopes achieving enhanced immunogenicity, the
CC      major benefit of the vaccine is that is safe and efficacious. AAB98391
CC      to AAB98477 represent polypeptide sequences used in the exemplification
CC      of the present invention.
SQ      Sequence 568 AA:

```

Query Match 99.0%; Score 2705; DB 22; Length 568;  
 Best Local Similarity 99.0%; Pred. No. 3.8e-247;  
 Matches 502; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

```

QY      1 MALWRESDNTVYLPPPSVARVYNTDDYVTRTSIFTHAGSSRLLTGPNPYFRVAPAGGNKQ 60
DB      62 MALWRESDNTVYLPPPSVARVYNTDDYVTRTSIFTHAGSSRLLTGPNPYFRVAPAGGNKQ 121
QY      61 DIPKVSAYQRYRFRVQLPDPNKFGLPDNSITYPEQRLVMAAGVEIGGQFLGVLGSLH 120
DB      122 DIPKVSAYQRYRFRVQLPDPNKFGLPDNSTIYPEQRLVMAAGVEIGGQFLGVLGSLH 181
QY      121 PFYNKLDPTRESSHAATSNSEDRDVSVDYKOTOLCIIGCAPAIGEHAKGNACKSRPL 180
DB      182 PFYNKLDPTRESSHAATSNSEDRDVSVDYKOTOLCIIGCAPAIGEHAKGNACKSRPL 241
QY      181 SOGDCPPLLEKNTVLEDGDMVDYGYGAMDFSTLDTKCEVPLDICSTICKYDPDYLOMSAD 240
DB      242 SOGDCPPLLEKNTVLEDGDMVDYGYGAMDFSTLDTKCEVPLDICSTICKYDPDYLOMSAD 301
QY      241 PYGDSMFCLRRRQFLARHFVNRAGVMTGTPQSLIKGTGNRASGSCVSPSPSGSIV 300
DB      302 PYGDSMFCLRRRQFLARHFVNRAGVMTGTPQSLIKGTGNRASGSCVSPSPSGSIV 361
QY      301 TSDSOLFNRPYWLHKAGHNGICWHNQLFVTVVDTTRSTNLTICASTOSPVGQYDARK 360
DB      362 TSDSOLFNRPYWLHKAGHNGICWHNQLFVTVVDTTRSTNLTICASTOSPVGQYDARK 421
QY      361 FKQYSRHEEYDLQITFQLCTITTLADVMSYIHSNMSILEDWNEGVPPPTSLVDYR 420
DB      422 FKQYSRHEEYDLQITFQLCTITTLADVMSYIHSNMSILEDWNEGVPPPTSLVDYR 481
QY      421 FVQSVAITCQKDAAPAEKNDPYDKLKFVWVDLKEKFSLDLDQYPLGRKFLVQAGLRKPT 480
DB      482 FVQSVAITCQKDAAPAEKNDPYDKLKFVWVDLKEKFSLDLDQYPLGRKFLVQAGLRKPT 541
QY      481 IGPKRKRSAPSATTTSSSKPAKRVRVRARK 507
DB      542 IGPKRKRSAPSATTTSSSKPAKRVRVRARK 568

```

```

RESULT 4
AAO18098
ID      AAO18098 standard; Protein; 500 AA.
AC      AAO18098;
DE      13-SEP-2002 (first entry)
DE      HPV type 18 L1 protein mutant encoded by gene l1delatc*.
KW      Human papillomavirus type 18; HPV; cancer; L protein; E protein;
KW      cytostatic; virucide; vaccine; lung; cervix; penis; vulva; anus.
OS      Human papillomavirus type 18.
OS      Synthetic.
PN      DE10059630-A1.
PD      06-JUN-2002.
PF      01-DEC-2000; 2000DE-1059630.
PR      01-DEC-2000; 2000DE-1059630.
PA      (MED1-) MEDIGENE AG.
PI      Mueller R, Nieland J, Gabelsberger J, Herbst R;
DR      WPI: 2002-520969/56.
DR      N-PSDB: AAL47486.
XX      Composition for treating and preventing tumors caused by human
XX      papilloma virus 18, contains fusion protein of papilloma virus L and E
XX      proteins, preferably truncated -
XX      Disclosure: Page 15-16; 64pp; German.

```

Query Match	98.6%	Score 2694	DB 23	Length 500
Best Local Similarity	99.6%	Pred. No. 3.4e-246		
Matches 498	Conservative 1	Mismatches 1	Indels 0	Gaps 0
Sequence	500 AA			
1	MAIMRPSNTYLYLPPPSVARYVNDYTRISIFHAASSRLTYGNYFRYPAGGKKQ	60		
1	MAIMRPSNTYLYLPPPSVARYVNDYTRISIFHAASSRLTYGNYFRYPAGGKKQ	60		
61	DIPKVSAYQYEVFRVQYLDPPNKGFLPDNSIYNPETORLVMAAGVEIGRQPLAGLSGH	120		
61	DIPKVSAYQYEVFRVQYLDPPNKGFLPDNSIYNPETORLVMAAGVEIGRQPLAGLSGH	120		
121	PFYVKLDTDESSHAATSNVSEEDYRDNVSDYKQYQCLILGAPALGEHMAKGTACKSRPL	180		
121	PFYVKLDTDESSHAATSNVSEEDYRDNVSDYKQYQCLILGAPALGEHMAKGTACKSRPL	180		
181	SGQDCPPELEKNTVLEDDMDVDTGAGADFSTLDPTKCEVPPLDICSICKYPPDYQMSAD	240		
181	SGQDCPPELEKNTVLEDDMDVDTGAGADFSTLDPTKCEVPPLDICSICKYPPDYQMSAD	240		
241	PYGDSMFECFLREDFARHFNNACTMGDTYPOSILYKGTGMRASPGSCVYSPSPSGSIV	300		
241	PYGDSMFECFLREDFARHFNNACTMGDTYPOSILYKGTGMRASPGSCVYSPSPSGSIV	300		
301	TSDSOLFNRKPYWLHKAQGNNGICGHNOLEFVTVVDTPSTNLTICASTQSPYPGQYDATK	360		
301	TSDSOLFNRKPYWLHKAQGNNGICGHNOLEFVTVVDTPSTNLTICASTQSPYPGQYDATK	360		
361	FKQYSRHYEEDYDLOFIFOLCTITTLADVMSYHSMNSSILEDWNGVPPPTSLVDYR	420		
361	FKQYSRHYEEDYDLOFIFOLCTITTLADVMSYHSMNSSILEDWNGVPPPTSLVDYR	420		
421	FVQSAITICQADAPAEKKDPYDKLFENNVDLKEKFSLDLQYPLGRREIVQAGLRKRP	480		
421	FVQSAITICQADAPAEKKDPYDKLFENNVDLKEKFSLDLQYPLGRREIVQAGLRKRP	480		
481	IGPRKRSAPSATTSKPAKR 500			
481	IGPRKRSAPSATTSKPAKR 500			

Query Match	Best Local Similarity	89.9%; Score 2457.5; DB 22; Length 536;
Matches 447; Conservative 29; Mismatches 31; Indels 3; Gaps 2;		
QY 1	MALMPPSDNFTVYLPSPVAVRVVNTDDYVTRTSIFHYAGSSRLLTVGNDFR-VPAAGGNNK 59	
DB 27	MALMPPSDSTVYLPSPVAVRVVNTDDYVTRTSIFHYAGSSRLLTVGNDFRVPVGAAGNNK 86	
QY 60	ODPKVSAVYQVRFVYVQVLDPPNKGLPNSIYNPPTQRLVMAACGVEFGROPGLGVGSG 119	
DB 87	QAVRKVSAVYQVRFVYVQVLDPPNKGLPNSIYNPPTQRLVMAACGVEFGROPGLGVGSG 146	
QY 120	HPFYNKLLDDESSHAATSNVSEEDVADVNDVYKQOTQLCLICAPAIIGEMHAKGTACKSRP 179	
DB 147	HPFYNKLLDDESSHAATSNVSEEDVADVNDVYKQOTQLCLICAPAIIGEMHAKGTACKSRP 206	
QY 180	LSQGDGCPLELAKNTYLEGDDVADVDSGYAMDSTLDDTFCBPRLDTCOSICTKPYDLQMSA 239	
DB 207	LQPGDGPPELAKNTYLEGDDVADVDSGYAMDSTLDDTFCBPRLDTCOSICTKPYDLQMSA 266	
QY 240	DPYGDSMFECFLRREQLFARHFMWNRAGTGVDPVQSLYTKGV--GMRAPSGSCVSPSPSG 297	
DB 267	DPYGDSMFECFLRREQLFARHFMWNRAGTGVDPVQSLYTKGV--GMRAPSGSCVSPSPSG 326	
QY 298	SVTNSDSOLENKKPYWLKRAQGHNNICWHNOLFVTVVDVTRSTNLTICASTQSPVPGQYD 357	
DB 327	SVTNSDSOLENKKPYWLKRAQGHNNICWHNOLFVTVVDVTRSTNLTICASTQSPVPGQYD 386	
QY 358	ATKRYOVSRAHVEEDLOEITPOLCTTLTADVMSYSHSNSSITLDMNNGVPRPTSLVD 417	
DB 387	PTKRHHYSRAHVEEDLOEITPOLCTTLTAEVMSYSHSNSSITLDMNNGVPRPTSLVD 446	
QY 418	TYREVQSAVTCQADDAFAENKDDPYDKLTKFNWVVDLKEKFSJLDDLOPYLGRFVLQAGRLR 477	

Db 447 TYFEVSVAVTCCKDPTPEKODPYDKLFWTVDLKEKSSDDQYPLGRKFLVQAGLR 506  
 QY 478 KPTIGPKRRSAPSATTSKPAKRVRRAR 507  
 507 RPTIGPKRRAPASTASRAKRVIRSK 536

## RESULT 6

AAB98454  
 ID AAB98454 standard; Protein; 534 AA.

AC AAB98454;

DT 22-AUG-2001 (first entry)

DE Human papillomavirus protein HPV56 L1.

KW Human papillomavirus; human leukocyte antigen; HLA; immune response;

KW HPV; epitope; T cell; identification; vaccine; infection; genital wart;

KW neoplastic growth; antiviral.

OS Human papillomavirus.

PN WO200141799-A1.

PD 14-JUN-2001.

PF 11-DEC-2000; 2000WO-US33549.

PR 10-DEC-1999; 9905-0172705.

PR 15-AUG-2000; 2000US-0641528.

PA (EPIM-) EPIMUNE INC.

PI Sette A, Sidney J, Southwood S, Chesnut R, Celis E, Grey HM;

DR WPI; 2001-381497/40.

PT An isolated human papilloma virus (HPV) epitope, useful in vaccines for

PS treating HPV infections -

XX Disclosure; Page 27; 756pp; English.

XX The present invention describes an isolated prepared human papillomavirus

CC (HPV) epitope (I). (I) has antiviral activity, and can be used in

CC vaccine production. Peptides and corresponding nucleic acid compositions

CC from the present invention are useful for stimulating an immune response

CC specifically in the treatment or prophylaxis of HPV infection, in persons

CC who have not manifested symptoms e.g. genital warts or neoplastic growth.

CC The peptides can also be used in a tetramer staining assay to assess

CC peripheral blood mononuclear cells for the presence of antigen-specific

CC CTLs following exposure to a pathogen or immunogen, and as reagents to

CC evaluate immune recall responses or evaluate the efficacy of a vaccine.

CC The vaccine compositions are useful for removing warts or treating HPV

CC infections. The epitopes for inclusion in an epitope-base vaccine may

CC be selected from conserved regions of viral or tumour-associated

CC antigens, which reduces the likelihood of escape mutants, also

CC immunosuppressive epitopes that may be present in whole antigens can be

XX Sequence 534 AA;

Query Match 67.7%; Score 1851; DB 22; Length 534;

Best Local Similarity 67.0%; Pred. No. 3,1e-166;

Matches 337; Conservative 71; Mismatches 89; Indels 6; Gaps 4;

QY 1 MALMRSDNTVYLPPEPSVAHVNTDDYVTRTSIFYHAGSSRLTLVGNPYRVPAGGNGK 60  
 Db 36 MATWRPSENNKYLPPTPVSKVATDSYKRTSIFYHAGSSRLLAHVHYSV-TKDNKT 94  
 QY 61 DIRKVASQYRVFRVQVLPDPNFKGLPONSITYNETQRLVWACAGVEIGRQPLGVLSGH 120  
 Db 95 NIPKVSAYQYRVFRVRLPDPNFKGLPNTNINPDQERLVWACAGVEIGRQPLGVLSGH 154  
 QY 121 PFYNKLDPTSSSHAATSNVSEEDVADNVSVYKQOTLCILCAPAIGBMAKTAACKSRPL 180  
 Db 155 PLFNRLDDPTSSNANANNVIEDSRDNI SVGKQOTLCIVGCTPAMGBHMKAVCKSTQV 214  
 QY 181 SQGDCPPELEKNTVLEDGDNVDTGYGAMDFSTLDQTKCEVPLDICOSICKYRPLYQMSAD 240  
 Db 215 TTGDCPPLALINPIEDGDMIDGFGAMDFKVLQESKAEPPLDIVGCTKPYDLKMSAD 274  
 QY 241 PYGDSMFPCRLRQDLFARHFNWNRAGTMDVTPQSLYIKGCGMARSPSCYSPSGSIV 300  
 Db 275 AYGDGMFYLRRQDLFARHFNWNRAGTMDVTPQSLYIKGCGMARSPSCYSPSGSIV 334  
 QY 301 TSDSOLEFKPYWLHKAQHNNGICWNLQFLVTVVDTFRSTNLTICASTQSPVGOYATK 360  
 Db 335 TSEKQLEFKPYWLQRAQGHNGICWNLQFLVTVVDTFRSTNLTICASTQSPVGOYATK 392  
 QY 361 FKOYSRHEEYDLOFQITLTLADVMSYIHSNSSLIEDNMFVPPPTSLVDTYR 420  
 Db 393 INQYLRHEVEYELQFVQOLKITLSAEVMAVYIHNMMANLLEDNMIGLSPVATSLDEK 452  
 QY 421 FVOSVAITCOKDAAPAKNKPYDKLKWVNDLKEKESLDDQYPLGKFLVQAGLRKPT 480  
 Db 453 YVRSTATTCOREOPPTPEKODPYDKLKWVNDLKEKESLDDQYPLGKFLVQAGLRKPT 512  
 QY 481 IG-PRKRSAPSATTSKPAKRV 502  
 Db 513 VATSKRSAP-ISTSTPAKRR 533

## RESULT 7

AAB98436  
 ID AAB98436 standard; Protein; 504 AA.

AC AAB98436;

DT 22-AUG-2001 (first entry)

DE Human papillomavirus protein HPV31 L1.

KW Human papillomavirus; human leukocyte antigen; HLA; immune response;

KW HPV; epitope; T cell; identification; vaccine; infection; genital wart;

KW neoplastic growth; antiviral.

OS Human papillomavirus.

PN WO200141799-A1.

PD 14-JUN-2001.

PF 11-DEC-2000; 2000WO-US33549.

PR 10-DEC-1999; 9905-0172705.

PR 15-AUG-2000; 2000US-0641528.

PA (EPIM-) EPIMUNE INC.

PI Sette A, Sidney J, Southwood S, Chesnut R, Celis E, Grey HM;

DR WPI; 2001-381497/40.

PT An isolated human papilloma virus (HPV) epitope, useful in vaccines for

PS treating HPV infections -

XX Disclosure; Page 23-24; 756pp; English.

CC The present invention describes an isolated prepared human papillomavirus  
 CC (HPV) epitope (I). (I) has antiviral activity, and can be used in  
 CC vaccine production. Peptides and corresponding nucleic acid compositions  
 CC from the present invention are useful for stimulating an immune response  
 CC to HPV by stimulating the production of CTL or HTL responses.  
 CC Specifically in the treatment or prophylaxis of HPV infection, in persons  
 CC who have not manifested symptoms e.g. genital warts or neoplastic growth.  
 CC The peptides can also be used in a tetramer staining assay to assess  
 CC peripheral blood mononuclear cells for the presence of antigen-specific  
 CC T cells following exposure to a pathogen or immunogen, and as reagents to  
 CC evaluate immune recall responses or evaluate the efficacy of a vaccine.  
 CC The vaccine compositions are useful for removing warts or treating HPV  
 CC infections. The epitopes for inclusion in an epitope-base vaccine may  
 CC be selected from conserved regions of viral or tumour-associated  
 CC antigens, which reduces the likelihood of escape mutants, also  
 CC immunosuppressive epitopes that may be present in whole antigens can be  
 CC avoided with the use of epitope-base vaccines. An additional advantage  
 CC is the ability to combine selected epitopes (CTL and HTL) and to modify  
 CC the composition of the epitopes achieving enhanced immunogenicity, the  
 CC major benefit of the vaccine is that is safe and efficacious. AAB98391  
 CC to AAB98477 represent polypeptide sequences used in the exemplification  
 CC of the present invention.

XX Sequence 504 AA:

Query Match 67.1%; Score 1834.5; DB 22; Length 504;  
 Best Local Similarity 65.1%; Pred. No. 1e-164;  
 Matches 329; Conservative 79; Mismatches 94; Indels 3; Gaps 3;

QY 1 MALNRPSNTYVYLPVPPSAVRYVNDYVTRTSISFYHAGSSRLTVGNFYPVAGGK 60  
 DB 1 MSIMRPSATYVYLPVPPSAVRYVNDYVTRTSISFYHAGSSRLTVGNFYPVAGGK 60  
 QY 61 -DIPKSAQYVFRVQLPDPNFKGLPNSIYVPEQRLVWACAGVEIGRQPLVGLSG 119  
 DB 61 IYVKEVSGIQYRFRVRLPDPNFKGFPDTSFYNPETQRLVWACVGLVGRQPLVGLSG 120  
 QY 120 HPFNKLDLTSSHAATSNVSEVDVNDVYKQQLCLICGAPAIIGHMAKGTCKSRP 179  
 DB 121 HPFNKLDLTSSHAATSNVSEVDVNDVYKQQLCLICGAPAIIGHMAKGTCKSRP 180  
 QY 180 LSGDCPPLLELKNYVLEGGDMVDYTGAMDEFTLDTCCEVLDICGSIKCYPDLOMSA 239  
 DB 181 ITPDCPPLLELKNYVLEGGDMVDYTGAMDEFTLDTCCEVLDICGSIKCYPDLOMSA 240  
 QY 240 PYGDSMFCLRRQLEFARHFWNRAGTMDVPOSLYIKGTGMRASPGSCVYSPSGSI 299  
 DB 241 EPGDTLFFYLRREQMEFARHFWNRAGTMDVPOSLYIKGTGMRASPGSCVYSPSGSI 300  
 QY 300 VTSOSLFNKRKYMLHKAGHNGICMHNOLFVYVDTFRSTNLTICASTOSPVRYQYDAT 359  
 DB 301 VTSOSLFNKRKYMLHKAGHNGICMHNOLFVYVDTFRSTNLTICASTOSPVRYQYDAT 359  
 QY 360 KFKOYSRHEVEYDLOFIFOLCTITLADVMSYIHSNMSISLDMNFGVPPPTSLVPTY 419  
 DB 360 NFKELRIRGEFEDLOFIFOLCTITLADVMSYIHSNMSISLDMNFGVPPPTSLVPTY 419  
 QY 420 RFVOSVAITCQKDAAPAKNDPYDKLFWNDLKEKESLDIQPLGKFTLVQAGLRKP 479  
 DB 420 RFVOSVAITCQKDAAPAKNDPYDKLFWNDLKEKESLDIQPLGKFTLVQAGLRKP 479  
 QY 480 TIGPRKRSAPSATSSRKARVRYR 504  
 DB 480 KFKAGKRSAPSATSSRKARVRYR 504

RESULT 8  
 AAG75800  
 ID AAG75800 standard; Protein: 505 AA.

XX AAG75800;  
 AC  
 XX 08-MAY-2003 (first entry)

XX Human papillomavirus 16 L1 protein, clone p14/16/11.  
 DE  
 XX HPV16: virus-like particle; VLP; infection; cancer; tumour; L1;  
 KW anogenital tract; uterine cervix; clone p14/16/11; virucide; vaccine.  
 KM  
 XX Human papillomavirus 16.  
 OS  
 XX US2002168372-A1.  
 PN  
 XX 14-NOV-2002.  
 PD  
 XX 29-SEP-1998; 98US-0162904.  
 PF  
 XX 16-JUL-1993; 93US-0092528.  
 PR 01-MAY-1996; 96US-0641570.  
 PR 27-JUN-1997; 97US-0884168.  
 XX (DURS/) DURST M.  
 PA (GISS/) GISSMANN L.  
 XX  
 XX Durst M, Gissmann L;  
 PI  
 XX WPI: 2003-275319/27.  
 DR N-PSDB; ABX11396.  
 XX  
 PT New DNA sequence encoding an L1 protein of a papillomavirus capable of  
 PT forming virus-like particles, useful for preparing a vaccine against  
 PT papillomavirus infection, particularly human papillomavirus (HPV)16  
 PT infection -  
 PS Disclosure: Fig 2; 17pp; English.

XX The invention discloses a DNA sequence encoding an L1 protein of a human  
 CC papillomavirus (HPV16) capable of forming virus-like particles (vlp).  
 CC Papillomaviruses infections have been linked to malignant cancer and  
 CC malignant tumours of the anogenital tract, particularly cancer of the  
 CC uterine cervix. Also disclosed is an antibody that is specifically  
 CC directed against the vlp or the L1 protein, a method for determining  
 CC anti-HPV16 virion antibodies in a sample and prophylaxis of  
 CC papillomavirus infections. The DNA sequence is useful for preparing a  
 CC vaccine against papillomavirus infection, particularly HPV16 infection.  
 CC The DNA is used to produce a VLP which can be used in a diagnostic kit  
 CC to determine anti-HPV16 virion antibodies in a sample. The VLP can also  
 CC be used to produce an L1 protein. The sequence presented is the HPV16 L1  
 CC protein, clone p14/16/11.  
 CC  
 XX  
 XX

SO Sequence 505 AA:

Query Match 67.0%; Score 1831; DB 24; Length 505;  
 Best Local Similarity 66.5%; Pred. No. 2.2e-164;  
 Matches 334; Conservative 68; Mismatches 96; Indels 4; Gaps 3;

QY 1 MALNRPSNTYVYLPVPPSAVRYVNDYVTRTSISFYHAGSSRLTVGNFYPVAGGK 60  
 DB 1 MSIMRPSATYVYLPVPPSAVRYVNDYVTRTSISFYHAGSSRLTVGNFYPVAGGK 60  
 QY 61 -DIPKSAQYVFRVQLPDPNFKGLPNSIYVPEQRLVWACAGVEIGRQPLVGLSG 119  
 DB 61 IYVKEVSGIQYRFRVRLPDPNFKGFPDTSFYNPETQRLVWACVGLVGRQPLVGLSG 120  
 QY 120 HPFNKLDLTSSHAATSNVSEVDVNDVYKQQLCLICGAPAIIGHMAKGTCKSRP 179  
 DB 121 HPFNKLDLTSSHAATSNVSEVDVNDVYKQQLCLICGAPAIIGHMAKGTCKSRP 180  
 QY 180 LSGDCPPLLELKNYVLEGGDMVDYTGAMDEFTLDTCCEVLDICGSIKCYPDLOMSA 239  
 DB 181 ITPDCPPLLELKNYVLEGGDMVDYTGAMDEFTLDTCCEVLDICGSIKCYPDLOMSA 240  
 QY 240 PYGDSMFCLRRQLEFARHFWNRAGTMDVPOSLYIKGTGMRASPGSCVYSPSGSI 299  
 DB 241 EPGDTLFFYLRREQMEFARHFWNRAGTMDVPOSLYIKGTGMRASPGSCVYSPSGSI 300  
 QY 300 VTSOSLFNKRKYMLHKAGHNGICMHNOLFVYVDTFRSTNLTICASTOSPVRYQYDAT 359  
 DB 301 VTSOSLFNKRKYMLHKAGHNGICMHNOLFVYVDTFRSTNLTICASTOSPVRYQYDAT 359  
 QY 360 KFKOYSRHEVEYDLOFIFOLCTITLADVMSYIHSNMSISLDMNFGVPPPTSLVPTY 419  
 DB 360 NFKELRIRGEFEDLOFIFOLCTITLADVMSYIHSNMSISLDMNFGVPPPTSLVPTY 419  
 QY 420 RFVOSVAITCQKDAAPAKNDPYDKLFWNDLKEKESLDIQPLGKFTLVQAGLRKP 479  
 DB 420 RFVOSVAITCQKDAAPAKNDPYDKLFWNDLKEKESLDIQPLGKFTLVQAGLRKP 479  
 QY 480 TIGPRKRSAPSATSSRKARVRYR 504  
 DB 480 KFKAGKRSAPSATSSRKARVRYR 504

```

QY 301 TSDSOLFNRPYWHLAKAGHNGICWGNOLFVTVVDTTRSTNLTICASTOSPVPGQDARK 360
DB 301 TSDAOLFNRPYWHLAKAGHNGICWGNOLFVTVVDTTRSTNLTICAAI-STSETTKNTN 359
QY 361 FKQYSHRVEEYDLOFIFOLCTITLTADVNSYIHSNMSSTLEDNFGVPPPTSLVDYR 420
DB 360 FKQYSHRVEEYDLOFIFOLCTITLTADVNTYIHSNMSSTLEDNFGVPPPTSLVDYR 419
QY 421 FVQSVAITCOKDAAPAKENKDPYDKLFWMVNDLKEKESLDDOYPLGRKFLVQAGLRKRP- 479
DB 420 FVTSOAIACOKHPPAPKEDPLKRYTFWEYNLKEKESADLDOPPLGRKFLVQAGLRKRP 479
QY 480 -TIGPKRRSAPSATSSKPAKR 500
DB 480 FTIGKRRK-ATPTTSTSTAKR 500

RESULT 9
AA08019
ID AAY08019 standard; Protein; 505 AA.
AC AAY08019;
XX
XX
DT 08-JUL-1999 (first entry)
XX
XX Human papilloma virus L1 protein.
XX
XX L1 protein; capsomer; virus; vaccine; infection; treatment; prevention;
XX cervical carcinoma; fusion protein; anti-capsid; antibody;
XX antigenicity.
XX
XX Human papilloma virus.
XX
XX WO0918220-A1.
XX
XX 15-APR-1999.
XX
XX PD 06-OCT-1998; 98WO-US20965.
XX
XX PR 06-OCT-1997; 97US-0944368.
XX
XX PA (LOYO) UNIV LOYOLA CHICAGO.
XX
XX PI Gissmann L, Mueller M;
XX
XX WPI; 1999-264026/22.
XX
XX N-PSDB; AAX37566.
XX
XX Human papilloma virus (HPV) L1 fusion protein capsomers, used in
XX vaccines against HPV infection
XX
XX Disclosure: Page 31-32; 48pp; English.
XX
XX This invention describes novel vaccines comprising a human papilloma
XX virus (HPV) capsomer, composed of a HPV L1 fusion protein or a HPV L1
XX truncated protein. The vaccines can prevent or treat human papilloma
XX virus infection. Such therapeutic vaccinations can be used for relief
XX against, e.g. cervical carcinoma. Construction of chimeric proteins
XX comprising amino acid residues from L1 protein, and e.g. E6 or E7
XX protein, which give rise to chimeric capsomers, combines prophylactic and
XX therapeutic functions of a vaccine. Capsomers can promote elimination of
XX persistently infected cells. Capsomers can also escape neutralization
XX by pre-existing anti-capsid antibodies and hence possess longer
XX circulating half-life as compared to chimeric virus-like particles.
XX The fusion protein, which forms the capsomer, provides increased
XX antigenicity.
XX
XX Sequence 505 AA;
XX
XX Query Match 66.8%; Score 1826; DB 20; Length 505;
XX Best Local Similarity 66.3%; Pred. No. 6 6e-164;
XX Matches 333; Conservative 68; Mismatches 97; Indels 4; Gaps 3;

```

```

QY 1 MALMRSDNTVYLPSPSVARVNTDDYVETSTIFYHAGSSRLTGVNPFYRPAAGGNKO 60
DB 1 MSALMRSEANVYLPSPSVARVNTDEYVARTNTIYHAGSRLLAVGHFPPIKPNMKI 60
QY 61 DIPKVSAYQYRVFVLPDPNKFGLPDNSIYNDETQRLVWACGVEIRGQPLVGLSGH 120
DB 61 LVPKVSGLQYRVFRIHLPDNKFGFPDTSFYNPDQRLVWACGVEIRGQPLVGLSGH 120
QY 121 PFNKTLDDESSHAASNSSEVDKVDVYKOTOLCIGCAPAIGEHAKGTACKSRPL 180
DB 121 PLINKLDDTENASAYANAGVDNRECI SMDYKOTOLCIGCKPPIGEBHGKSPCTNVAV 180
QY 181 SGGDCPPELEKNTVLEDDGMVDVGYGAMPFSTLDTKCEVPLDICO SICKPYLQMSAD 240
DB 181 NPDCPPELEKNTVLEDDGMVDVGYGAMPFSTLDTKCEVPLDICO SICKPYLQMSAD 240
QY 241 PYGDSMEFCLRRQLFARHFNRAAGTMDTVPOSLYIKGTGRASPSCVYSPSGSIY 300
DB 241 PYGDSMEFCLRRQLFARHFNRAAGTMDTVPOSLYIKGTGRASPSCVYSPSGSIY 300
QY 301 TSDSOLFNRPYWHLAKAGHNGICWGNOLFVTVVDTTRSTNLTICASTOSPVPGQDARK 360
DB 301 TSDAOLFNRPYWHLAKAGHNGICWGNOLFVTVVDTTRSTNLTICAAI-STSETTKNTN 359
QY 361 FKQYSHRVEEYDLOFIFOLCTITLTADVNSYIHSNMSSTLEDNFGVPPPTSLVDYR 420
DB 360 FKQYSHRVEEYDLOFIFOLCTITLTADVNTYIHSNMSSTLEDNFGVPPPTSLVDYR 419
QY 421 FVQSVAITCOKDAAPAKENKDPYDKLFWMVNDLKEKESLDDOYPLGRKFLVQAGLRKRP- 479
DB 420 FVTSOAIACOKHPPAPKEDPLKRYTFWEYNLKEKESADLDOPPLGRKFLVQAGLRKRP 479
QY 480 -TIGPKRRSAPSATSSKPAKR 500
DB 480 FTIGKRRK-ATPTTSTSTAKR 500

RESULT 10
AA57720
ID AAY57720 standard; Protein; 505 AA.
AC AAY57720;
XX
XX
XX 14-MAR-2000 (first entry)
XX
XX DE Human papillomavirus L1 fusion protein SEQ ID NO:2.
XX
XX KW Human papillomavirus; HPV; L1 fusion protein; vaccine; cytostatic;
XX viral capsomere; virucide; dermatologically; malignant tumour formation;
XX condylomata acuminata.
XX
XX OS Human papillomavirus.
XX
XX Synthetic.
XX
XX CA2229955-A1.
XX
XX PD 20-AUG-1999.
XX
XX PE 20-FEB-1998; 98CA-2229955.
XX
XX PR 20-FEB-1998; 98CA-2229955.
XX
XX PA (MED-) MEDIGENE GMBH.
XX
XX PI Burger A, Hallek M;
XX
XX WPI; 2000-063092/06.
XX
XX N-PSDB; AA248174.
XX
XX Fusion proteins comprising papillomavirus specific proteins useful for
XX vaccinating against malignant tumors of the anogenital tract such as

```





DB 480 FTLGKRR-ATPTSTSTTAKR 500

RESULT 12  
AAB77479  
ID AAB77479 standard; Protein; 505 AA.  
XX  
XX AAB77479;  
XX  
XX 22-JUL-2002 (first entry)  
XX  
XX HPV16-L1 2.  
DE HPV16-L1 2.  
XX Human papilloma virus; HPV; HPV16-L1; HPV16-L2; vaccine; cervical cancer;  
XX HPV16-L1; cytostatic; virucide.  
XX  
XX Human papillomavirus.  
XX  
XX WO200238769-A2.  
XX  
XX 16-MAY-2002.  
XX  
XX 19-SEP-2001; 2001WO-DE03618.  
XX  
XX 09-NOV-2000; 2000DE-1055545.  
XX  
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
XX (IPKP-) IPK INST PFLANZENGENETIK & KULTURPFLANZE.  
XX  
XX Mueller M, Leder C, Kleinschmidt J, Sonnwald U, Biemelt S;  
XX WPI; 2002-426950/45.  
XX N-PSDB; ABL58977, ABL58982.  
XX  
XX New DNA sequences encoding human papilloma virus L1 or L2 protein,  
XX useful in vaccines, are optimized for high-level expression in  
XX eukaryotic cells .  
XX  
XX Claim 1; Fig 6; 39pp; German.  
XX  
XX The invention relates to DNA sequences (I) that encode human papilloma  
XX virus (HPV) L1 protein (ABL58976, ABL58977) or L2 protein (ABL58978)  
XX or proteins with the biological activity of L1 and L2. Expression vectors  
XX containing (I) or a similar sequence encoding an L1/L2 fusion protein  
XX (ABL58979-ABL58981) and the proteins encoded by them (AAB77478-AAB77483),  
XX are useful in vaccines, especially for control of cervical cancer. (I)  
XX are also useful for recombinant production of L1 and L2 proteins. (I) are  
XX optimised for codon usage in eukaryotic cells and provide high yields of  
XX L1/L2 or their fusions, without the use of viral vectors.  
XX  
XX Sequence 505 AA;  
SQ

Query Match 66.8%; Score 1826; DB 23; Length 505;  
Best Local Similarity 66.3%; Pred. No. 6,6e-164;  
Matches 333; Conservative 68; Mismatches 97; Indels 4; Gaps 3;

QY 1 MALMRPSDNTVYLLPPPSVAVVNTDDYVTRTSIFHYHAGSSRLITVGNPFYRPVAGGKNO 60  
DB 1 MSIMLPSEATVYLLPPVPSKVSTDEYVARKNIYHAGSRLLAAGHPFPFKKPNMKI 60  
QY 61 DIPKVASOYRVFVQVLPDPNFKGLPDSNIYVETQRLVWACAGVEIGRGQPLGVLSGH 120  
DB 61 IVPKVSGLQYRVFRIHLDPNKFEPDYSFYNDPQRLVWACGVGVGRGQPLGVLSGH 120  
QY 121 PYYNKLDOTSESSHAATSNSEDEVNDVSVYKOTQCLICGCAPIGEHNAKGTACKSRPL 180  
DB 121 PLNKLIDTENSAAVAANAGVNRRECISMDYKOTQCLICGKRPPIEGHNGKSPCTNVAV 180  
QY 181 SOGDPEPLELKNVTLEDGDMVDYGYGAMDFSLQDPKCEVPLDICOISICKYDPYLOMSAD 240  
DB 181 NGDDEPPELEINTVYQDGMVDYTGCAAMPFTLLQANKSEVPLDICTICKYDPYIMWSE 240  
QY 241 PYGDSMFCLRRQQLFARHFVNRAGTMDYVPOSILYIKGTGMRASPSGCVSPSPSGSIV 300

DB 241 PYGDSMFCLRRQQLFARHFVNRAGTMDYVPOSILYIKGTGMRASPSGCVSPSPSGSMV 300  
QY 301 TSDSOLEFNKPYMLHKQGNHNGICWNOLEFVYVVDTRSTNLFICASTQSPVPGYDARK 360  
DB 301 TSDAQLEFNKPYMLQRAQHNHNGICWNOLEFVYVVDTRSTNLSLCAI-STSEYTKKN 359  
QY 361 FKQYSHVEEYDQLEFQCTITLTDVAVSYIHSNSSLDEDMNFEVPPPTSLVDTYR 420  
DB 360 FKEYLHNGEYDQLEFQCTITLTDVAVSYIHSNSSLDEDMNFEQPPPGCTLEDYTR 419  
QY 421 FVOSVAITCQKDAAPENKDPYDKLFVAVDLKEKFSLDQYPLGKFLVQAGLRKP- 479  
DB 420 FVTSQAIACQKHPPAPKEDPPLKTYFWEVNLKEKFSADLDQFPILGRKFLAQLKARK 479  
QY 480 -TIGPKRRASPSATSSKPAKR 500  
DB 480 FTLGKRR-ATPTSTSTTAKR 500

RESULT 13  
AAB98422  
ID AAB98422 standard; Protein; 531 AA.  
XX  
XX AAB98422;  
XX  
XX 22-AUG-2001 (first entry)  
XX  
XX Human papillomavirus protein HPV16 L1.  
XX  
XX Human papillomavirus; human leukocyte antigen; HLA; immune response;  
XX HPV; epitope; T cell; identification; vaccine; infection; genital wart;  
XX neoplastic growth; antiviral.  
XX  
XX Human papillomavirus.  
XX  
XX WO200141799-A1.  
XX  
XX 14-JUN-2001.  
XX  
XX 11-DEC-2000; 2000WO-US33549.  
XX  
XX 10-DEC-1999; 99US-0172705.  
XX PR 15-AUG-2000; 2000US-0641528.  
XX  
XX (EPIIM-) EPIIMUNE INC.  
XX  
XX Sette A, Sidney J, Southwood S, Chesnut R, Celis E, Grey HM;  
XX WPI; 2001-381497/40.  
XX  
XX An isolated human papilloma virus (HPV) epitope, useful in vaccines for  
XX treating HPV infections -  
XX  
XX Disclosure; Page 21; 756pp; English.  
XX

The present invention describes an isolated prepared human papillomavirus (HPV) epitope (I). (I) has antiviral activity, and can be used in vaccine production. Peptides and corresponding nucleic acid compositions from the present invention are useful for stimulating an immune response to HPV by stimulating the production of CTL or HTL responses, specifically in the treatment or prophylaxis of HPV infection. In persons who have not manifested symptoms e.g. genital warts or neoplastic growth. The peptides can also be used in a tetramer staining assay to assess peripheral blood mononuclear cells for the presence of antigen-specific CTLs following exposure to a pathogen or immunogen, and as reagents to evaluate immune recall responses or evaluate the efficacy of a vaccine. The vaccine compositions are useful for removing warts or treating HPV infections. The epitopes for inclusion in an epitope-base vaccine may be selected from conserved regions of viral or tumour-associated antigens, which reduces the likelihood of escape mutants, also immunosuppressive epitopes that may be present in whole antigens can be avoided with the use of epitope-base vaccines. An additional advantage

CC is the ability to combine selected epitopes (CTL and HTL) and to modify  
CC the composition of the epitopes achieving enhanced immunogenicity, the  
CC major benefit of the vaccine is that is safe and efficacious. AAB98391  
CC to AAB98477 represent polypeptide sequences used in the exemplification  
CC of the present invention.

XX Sequence 531 AA;

Query Match 66.8%; Score 1826; DB 22; Length 531;

Best Local Similarity 66.3%; Pred. No. 7,1e-164;  
Matches 333; Conservative 68; Mismatches 97; Indels 4; Gaps 3;

```
QY 1 MAMRPSDNTVYLPSPSVARVNTDDYVTSIFHYAGSSRLTVGNPFRVAGGNGKQ 60
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 27 MSIMLSEAVTVLPVPVSKVSTDEVARTNITYHAAGTSRLAVAGHPPIKPNNNKI 86
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 DIPKVSAYQYRVFVNOVDPNKKFGLPDINSIYNPETORLVACAGVEVGRGQPLGVLSGH 120
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 87 LVRKVSGLQYRVRIHLPDPNKKFGPPTSTYNDPTORLVACVGEVGRGQPLGVLSGH 146
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 PFYNNKLDDESSHAATSNVEDVDVNYVDYKOTQCLICCAPAIGEHMAKGTACKSRPL 180
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 147 PLINKLDDTENAASAYANAGVDRNRECSIMDYKOTQCLICCKPPIGEHWKGSPTCNVAV 206
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 SGGDCPPELEKNTVLEDDGVDTGCGAMDESTLDPTKCEVPLDICSICKPYDLOMSAD 240
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 207 NPDCPPELELINTVLDGDMVDYGFAMDTFTVLQANKSEVPLDICSICKPYDIKWVSE 266
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 PYGDSMEFCLRRQOLFARHFWNRAGTMDVTPQSLYIKGTGMRASPGSCVSPSGSIV 300
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 267 PYGDSLEFYLRRQOMFRLHFNRAAGVGENVPDLYIKGSGSRANLASSYTFPPSSMY 326
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 TSSQLEFNKPYWTHKAQGNNGICWHNQLFVTVVDTTRSTNLTICASTQSPVPGQYDAK 360
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 327 TSDAQLEFNKPYWTHKAQGNNGICWHNQLFVTVVDTTRSTNLTICASTQSPVPGQYDAK 385
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 FKQYSRHVEYDLOFTFOLCTITLADVMSYIHSNNSIIEEDNFGVPPPTSLVDTYR 420
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 386 FKQYSRHVEYDLOFTFOLCTITLADVMSYIHSNNSIIEEDNFGVPPPTSLVDTYR 445
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 FVQSVAITCQKDAPEAKNDPYDKLFKFWNDLKEKESLDIOYPLGRKFLVQAGLRKRP- 479
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 446 FVTSQALACQKHTPPAKEDPLKRYFWEVNLKEKESADLDQPLGRKFLVQAGLRKRP 505
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 480 -TIGPKRRSAPSAATSSKPAKR 500
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 506 FTLGKRR-ARPTTSSSTSTAKR 526
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

## RESULT 14

ABG75799 standard; Protein; 505 AA.

```
XX ID ABG75799 standard; Protein; 505 AA.
XX AC ABG75799;
XX DT 08-MAY-2003 (first entry)
XX DE Human papillomavirus 16 L1 protein, clone P114/16/2.
XX KW HPV16; Virus-like particle; VLP; Infection; cancer; tumour; LI;
XX MM anogenital tract; uterine cervix; clone P114/16/2; virucide; vaccine.
XX OS Human papillomavirus 16.
XX PN US2002168372-A1.
XX PD 14-NOV-2002.
XX PF 29-SEP-1998; 9805-0162904.
XX PR 16-JUL-1993; 9305-0092528.
XX PR 01-JAN-1996; 9605-0641570.
XX PR 27-JUN-1997; 9705-0884168.
```

XX (DURS/) DURST M.  
XX PA (GISS/) GISSMANN L.  
XX PI Durst M, Gissmann L;  
XX WPI; 2003-275319/27.  
XX DR N-PSDB; ABX11395.  
XX PT New DNA sequence encoding an L1 protein of a papillomavirus capable of  
XX forming virus-like particles, useful for preparing a vaccine against  
XX papillomavirus infection, particularly human papillomavirus (HPV)16  
XX infection.  
XX Disclosure; Fig 1; 17pp; English.

CC The invention discloses a DNA sequence encoding an L1 protein of a human  
CC papillomavirus (HPV16) capable of forming virus-like particles (VLPs).  
CC Papillomaviruses infections have been linked to malignant cancer and  
CC malignant tumours of the anogenital tract, particularly cancer of the  
CC uterine cervix. Also disclosed is an antibody that is specifically  
CC directed against the VLP or the L1 protein, a method for determining  
CC anti-HPV16 virion antibodies in a sample and prophylaxis of  
CC papillomavirus infections. The DNA sequence is useful for preparing a  
CC vaccine against papillomavirus infection, particularly HPV16 infection.  
CC The DNA is used to produce a VLP which can be used in a diagnostic kit  
CC to determine anti-HPV16 virion antibodies in a sample. The VLP can also  
CC be used for the prophylaxis of papillomavirus infections and the DNA  
CC used to produce an L1 protein. The sequence presented is the HPV16 L1  
CC protein, clone P114/16/2.

XX Sequence 505 AA;

Query Match 66.8%; Score 1825; DB 24; Length 505;

Best Local Similarity 66.1%; Pred. No. 8,2e-164;  
Matches 332; Conservative 69; Mismatches 97; Indels 4; Gaps 3;

```
QY 1 MAMRPSDNTVYLPSPSVARVNTDDYVTSIFHYAGSSRLTVGNPFRVAGGNGKQ 60
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1 MSIMLSEAVTVLPVPVSKVSTDEVARTNITYHAAGTSRLAVAGHPPIKPNNNKI 60
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 DIPKVSAYQYRVFVNOVDPNKKFGLPDINSIYNPETORLVACAGVEVGRGQPLGVLSGH 120
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 61 LVRKVSGLQYRVRIHLPDPNKKFGPPTSTYNDPTORLVACVGEVGRGQPLGVLSGH 120
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 121 PFYNNKLDDESSHAATSNVEDVDVNYVDYKOTQCLICCAPAIGEHMAKGTACKSRPL 180
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 121 PLINKLDDTENAASAYANAGVDRNRECSIMDYKOTQCLICCKPPIGEHWKGSPTCNVAV 180
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 181 SGGDCPPELEKNTVLEDDGVDTGCGAMDESTLDPTKCEVPLDICSICKPYDLOMSAD 240
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 181 NPDCPPELELINTVLDGDMVDYGFAMDTFTVLQANKSEVPLDICSICKPYDIKWVSE 240
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 241 PYGDSMEFCLRRQOLFARHFWNRAGTMDVTPQSLYIKGTGMRASPGSCVSPSGSIV 300
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 241 PYGDSLEFYLRRQOMFRLHFNRAAGVGENVPDLYIKGSGSRANLASSYTFPPSSMY 300
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 301 TSSQLEFNKPYWTHKAQGNNGICWHNQLFVTVVDTTRSTNLTICASTQSPVPGQYDAK 360
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 301 TSDAQLEFNKPYWTHKAQGNNGICWHNQLFVTVVDTTRSTNLTICASTQSPVPGQYDAK 359
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 361 FKQYSRHVEYDLOFTFOLCTITLADVMSYIHSNNSIIEEDNFGVPPPTSLVDTYR 420
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 360 FKQYSRHVEYDLOFTFOLCTITLADVMSYIHSNNSIIEEDNFGVPPPTSLVDTYR 419
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 421 FVQSVAITCQKDAPEAKNDPYDKLFKFWNDLKEKESLDIOYPLGRKFLVQAGLRKRP- 479
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 420 FVTSQALACQKHTPPAKEDPLKRYFWEVNLKEKESADLDQPLGRKFLVQAGLRKRP 479
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 480 -TIGPKRRSAPSAATSSKPAKR 500
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 480 FTLGKRR-ARPTTSSSTSTAKR 500
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

RESULT 15  
 AAR38807  
 ID AAR38807 standard; Protein; 505 AA.  
 AC AAR38807;  
 XX 25-MAR-2003 (updated)  
 DT 17-DEC-2001 (updated)  
 DT 21-JAN-1994 (first entry)  
 XX  
 DE BPV1 L1.  
 XX  
 KW L1; capsid protein; bovine; human; papillomavirus; baculovirus;  
 KW transfer vector; promoter; capsid protein; transformation; vaccine;  
 KW neutralising antibody; vertebrate.  
 XX  
 OS Bovine papillomavirus.  
 XX  
 PN USN8032869-N.  
 PD 15-JUL-1993.  
 XX  
 PF 16-MAR-1993; 93US-0032869.  
 XX  
 PR 03-SEP-1992; 92US-0941371.  
 PR 16-MAR-1993; 93US-0032869.  
 XX  
 PA (USSH ) US DEPT HEALTH & HUMAN SERVICE.  
 XX  
 PI Kirnbauer R, Lowy DR, Schiller JT;  
 DR WPI; 1993-249995/31.  
 DR N-PSDB; AAQ47166.  
 XX  
 PT Recombinant papilloma virus capsid proteins - for vaccines  
 PT against papilloma virus and for diagnosis of virus infection  
 XX  
 PS Example 1; Page 32-34; 45pp; English.  
 XX  
 CC The sequences given in AAR38807-08 represent the L1 capsid proteins from  
 CC bovine and human papillomavirus respectively. The DNA encoding these  
 CC sequences may be inserted into a baculovirus transfer vector and  
 CC operatively expressed by a promoter of the vector, and the capsid  
 CC protein produced by transformed cells. These capsid proteins may be  
 CC used in vaccines to induce high-titre neutralising antibody response  
 CC in vertebrates.  
 CC (Note: Revised entry submitted to correct the patent number format of  
 CC US Government-owned NTIS applications to prevent clashes with ongoing US  
 CC granted patent numbers. For further information please visit the Derwent  
 CC web site at [www.derwent.com/dwpi/updates/ntis-us.html](http://www.derwent.com/dwpi/updates/ntis-us.html).)  
 CC  
 CC (Updated on 25-MAR-2003 to correct PF field.)  
 CC  
 SQ Sequence 505 AA;  
 QY  
 Query Match 66.7%; Score 1824; DB 14; Length 505;  
 Best Local Similarity 66.3%; Pred. No. 1e-163;  
 Matches 333; Conservative 68; Mismatches 97; Indels 4; Gaps 3;  
 Db 1 MALMRSPDMVTYLPSPSVARVYVTDYVTRTSIFYHAGSSRLITVGNPFRVAGGKNQ 60  
 1 MSIMLPSEATVYLPVPSKVVSTDEYVAFTNITYHAGTSRLAVGHPYPIKPPNNNKI 60  
 QY 61 DIPKVSAYQYRVFRVQLPDPNKGFLPDNSIYNPETORLWACAGVEIGRGQPLGVLGSH 120  
 61 IVPKVSGLQYRVFRIRILPDPNKGFPDTSFYNDIQLWACAGVEIGRGQPLGVLGSH 120  
 Db 61 IVPKVSGLQYRVFRIRILPDPNKGFPDTSFYNDIQLWACAGVEIGRGQPLGVLGSH 120  
 QY 121 PFYNNKLDTESSHAATSNVSEVDYNDVSDYKOTQLCIGCAPAIGEHMAKGTACKSRPL 180  
 121 PLTNKLDTEENASAVYANAGVDNRRECISMDYKOTQLCIGCKPPIGEHMGKSGSPCTNVAV 180  
 Db 121 PLTNKLDTEENASAVYANAGVDNRRECISMDYKOTQLCIGCKPPIGEHMGKSGSPCTNVAV 180  
 QY 181 SGGDCPPLELKNVYLDGDMVDTGCMDFSTLQDTKCFVPLDICSICKYPRYLDMSAD 240  
 181 SGGDCPPLELKNVYLDGDMVDTGCMDFSTLQDTKCFVPLDICSICKYPRYLDMSAD 240

Db 181 NPQDCPPLELKNVYLDGDMVDTGCMDFSTLQANKSEVPDICTSICKYPRYIMVSE 240  
 QY 241 PYGDSMEFCLRRQLEFARHFWNRAGTWDVPOSLYIKGTGMRASPGCYSPSPSGSIY 300  
 241 PYGDSLEFYLRRQGMFVRHLFNAGTYGGEVPPDLYIKSGSGSTANLASSNYFFTPSGSMY 300  
 Db 241 PYGDSLEFYLRRQGMFVRHLFNAGTYGGEVPPDLYIKSGSGSTANLASSNYFFTPSGSMY 300  
 QY 301 TSDSQLEFNKPYWLHKAQGNNGICWNNQLEFVAVDTRSTNLTICASTQSPVPGQYDARK 360  
 301 TSDAQIFNKPWYLQRAQGNNGICWNNQLEFVAVDTRSTNLSICAAI-STSEITYKNTN 359  
 Db 301 TSDAQIFNKPWYLQRAQGNNGICWNNQLEFVAVDTRSTNLSICAAI-STSEITYKNTN 359  
 QY 361 FKQSRHVEEYDLOFLOCTITLTDVMSYHSMNSTLEDNMFVPPPTSLVDYR 420  
 361 FKELRRGEEDLOFLOCTITLTDVMTYHISMNSTILEDNMFGLQPPGGTLEDTYR 419  
 Db 361 FKELRRGEEDLOFLOCTITLTDVMTYHISMNSTILEDNMFGLQPPGGTLEDTYR 419  
 QY 421 FVQSVAITCQDAAPAEKNDPYDLKFWNDLKEKESLDDQYPLGRKFLVQGLRRKP- 479  
 421 FVTSQAIACQKHPPAPKEDPLAKKTYFEVNLKESADLDQPLGRKFLVQGLRKP 479  
 Db 421 FVTSQAIACQKHPPAPKEDPLAKKTYFEVNLKESADLDQPLGRKFLVQGLRKP 479  
 QY 480 -TIGPKRRASAPSATTSKPAKR 500  
 480 FTLGKRR-ATPTSSTSTAKR 500  
 Db 480 FTLGKRR-ATPTSSTSTAKR 500

Search completed: October 17, 2003, 10:59:07  
 Job time : 64.8512 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2003, 10:47:38 ; Search time 15.7128 Seconds  
(without alignments)  
1517.395 Million cell updates/sec

Title: US-08-913-644-2

Perfect score: 2733  
Sequence: 1 MALMRPSDNTVYLPPPSVAR.....APSATTSPKARRVRARRK 507

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2705	99.0	568	VLI_HPV18	P06794 human papill
2	2450	89.6	539	VLI_HPV45	P36741 human papill
3	2179.5	79.7	504	VLI_HPV70	P50793 human papill
4	2155	78.9	505	VLI_HPVME	P27964 human papill
5	2151	78.7	505	VLI_HPV39	P24838 human papill
6	2141	78.3	505	VLI_HPV68	P54669 human papill
7	1956	71.6	503	VLI_HPV29	P50792 human papill
8	1917.5	70.2	532	VLI_HPV03	P36731 human papill
9	1912.5	70.0	502	VLI_HPV28	P50791 human papill
10	1909	69.8	531	VLI_HPV10	P36732 human papill
11	1879	68.8	508	VLI_HPV30	P02515 human papill
12	1861.5	68.1	503	VLI_HPV26	P36735 human papill
13	1855.5	67.9	505	VLI_HPV61	P50892 human papill
14	1855	67.9	499	VLI_HPV53	P05113 human papill
15	1855	67.9	501	VLI_HPV1	P22163 rhesus papill
16	1851	67.7	534	VLI_HPV56	P36743 human papill
17	1837	67.2	497	VLI_HPV54	P50819 human papill
18	1834.5	67.1	504	VLI_HPV31	P17388 human papill
19	1822	66.7	502	VLI_HPV66	P080961 human papill
20	1820.5	66.6	502	VLI_HPV35	P27222 human papill
21	1804.5	66.0	504	VLI_HPV51	P26536 human papill
22	1804.5	66.0	524	VLI_HPV58	P26535 human papill
23	1801.5	65.9	529	VLI_HPV52	P05138 human papill
24	1801	65.9	531	VLI_HPV16	P03101 human papill
25	1791	65.5	499	VLI_HPV33	P06416 human papill
26	1789.5	65.5	510	VLI_HPV2A	P25466 human papill
27	1784.5	65.3	510	VLI_HPV2A	P22162 human papill
28	1782	65.2	502	VLI_PCPV1	P02274 pygmy chipm
29	1781	65.2	594	VLI_HPV27	P36736 human papill
30	1771	64.8	505	VLI_HPV07	P05136 human papill
31	1768.5	64.7	500	VLI_HPV6A	P03100 human papill
32	1766	64.6	505	VLI_HPV40	P36740 human papill
33	1763	64.5	528	VLI_HPV34	P36738 human papill

34	1762	64.5	503	VLI_HPV32	P36737 human papill
35	1759	64.4	499	VLI_HPV13	P02273 human papill
36	1757	64.3	502	VLI_HPV42	P27233 human papill
37	1751.5	64.1	501	VLI_HPV55	P50820 human papill
38	1745	63.8	500	VLI_HPV44	P50816 human papill
39	1733	63.4	501	VLI_HPV11	P06417 human papill
40	1427	52.2	514	VLI_HPV08	P02050 human papill
41	1426.5	52.2	546	VLI_HPV19	P06417 human papill
42	1426	52.2	516	VLI_HPV20	P50786 human papill
43	1425.5	52.2	517	VLI_HPV12	P50786 human papill
44	1422.5	52.0	517	VLI_HPV25	P02051 human papill
45	1421	52.0	510	VLI_HPV38	P50814 human papill

## ALIGNMENTS

```

RESULT 1
VLI_HPV18          STANDARD;          PRT;          568 AA.
ID VLI_HPV18
AC P06794; 084270;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Major capsid protein l1.
GN l1.
OS Human papillomavirus type 18.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10582;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=87283882; PubMed=3039146;
RA Cole S.T., Danos O.;
RT "Nucleotide sequence and comparative analysis of the human
RT papillomavirus type 18 genome. Phylogeny of papillomaviruses and
RT repeated structure of the E6 and E7 gene products.";
RL J. Mol. Biol. 193:599-608(1987).
RN [2]
RP SEQUENCE OF 363-406 FROM N.A.
RA MEDLINE=92407963; PubMed=1326639;
RA Chan S.Y., Bernard H.U., Ong C.K., Chan S.P., Birgit H., Delius H.;
RT "Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and
RT variants: a showcase for the molecular evolution of DNA viruses.";
RL J. Virol. 66:5714-5725(1992).
RN [3]
RP SEQUENCE OF 501-568 FROM N.A.
RA MEDLINE=87061231; PubMed=3023691;
RA Thierry F., Heard J.M., Dartmann K., Yaniv M.;
RT "Characterization of a transcriptional promoter of human
RT papillomavirus 18 and modulation of its expression by simian virus 40
RT and adenovirus early antigens.";
RL J. Virol. 61:134-142(1987).
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: X05015; CA28671.1; -
DR EMBL: M96287; AAA47026.1; -
DR EMBL: M14710; AAA65508.1; -
DR EMBL: A06329; CA00545.1; -
DR PIR: A26251; P1WL18.
DR InterPro: IPR002210; PV_capsid_l1.
DR Pfam: PF00500; late_protein_l1.
DR PRINTS: PR00865; HPVcapsid_l1.
DR ProDom: PD000544; PV_capsid_l1.
DR Coat protein: Late protein.
KW CONFLICT 384 384 V -> I (IN REF. 2).

```

SQ SEQUENCE 568 AA; 63623 MW; BB2BF6361177FC10 CRC64;  
 Query Match 99.0%; Score 2705; DB 1; Length 568;  
 Best Local Similarity 99.0%; Pred. No. 8.7e-227;  
 Matches 502; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MALMRPSNTYLLPPPSVARYVNDYTRKSIYFHAGSSRLTYGNFYFRYPAGGKQ 60  
 |||||||  
 DB 62 MALMRPSNTYLLPPPSVARYVNDYTRKSIYFHAGSSRLTYGNFYFRYPAGGKQ 121  
 |||||||  
 QY 61 DIPKVSAYQYRVFVQVLPDPNKKFGLPDNSIYNPEQRLVWACAGVEIGROPGLGSLG 120  
 |||||||  
 DB 122 DIPKVSAYQYRVFVQVLPDPNKKFGLPDNSIYNPEQRLVWACAGVEIGROPGLGSLG 181  
 |||||||  
 QY 121 PFYFKLDDTESSHAATSNVSEVDVNDVSDYKQVQLCICGAPALGEHMAKGTACKSRPL 180  
 |||||||  
 DB 182 PFYFKLDDTESSHAATSNVSEVDVNDVSDYKQVQLCICGAPALGEHMAKGTACKSRPL 241  
 |||||||  
 QY 181 SGGDCPRLKNTYLEDDGMVDYGYGAMDESTLDQTKCEVPLDQICQICKYPPYLQMSA 240  
 |||||||  
 DB 242 SGGDCPRLKNTYLEDDGMVDYGYGAMDESTLDQTKCEVPLDQICQICKYPPYLQMSA 301  
 |||||||  
 QY 241 PYGDSMFECRLRQDLFAHFNNRAGTMDYVPOSILYIKGTGMASPGSCYSPSPSGSIV 300  
 |||||||  
 DB 302 PYGDSMFECRLRQDLFAHFNNRAGTMDYVPOSILYIKGTGMASPGSCYSPSPSGSIV 361  
 |||||||  
 QY 301 TSDSOLFENKPYWHLKAGCHNNGICWHNOLFVTVVDTTRSTNLITCASTQSPVPGQYDARK 360  
 |||||||  
 DB 362 TSDSOLFENKPYWHLKAGCHNNGICWHNOLFVTVVDTTRSTNLITCASTQSPVPGQYDARK 421  
 |||||||  
 QY 361 FKQSRHVEEYDLOFIFOLCTITLADYMSYIHSNSSLIEDMNGVPPPTSLVDYR 420  
 |||||||  
 DB 422 FKQSRHVEEYDLOFIFOLCTITLADYMSYIHSNSSLIEDMNGVPPPTSLVDYR 481  
 |||||||  
 QY 421 FVQSVATTCQKDAAPAEKNDYDKLKFNNVDLKEKESLDLDQYPLGRKFLVQAGLRKPT 480  
 |||||||  
 DB 482 FVQSVATTCQKDAAPAEKNDYDKLKFNNVDLKEKESLDLDQYPLGRKFLVQAGLRKPT 541  
 |||||||  
 QY 481 IGPKRKRSAPATSSKPAKRVVRARK 507  
 |||||||  
 DB 542 IGPKRKRSAPATSSKPAKRVVRARK 568  
 |||||||

RESULT 2  
 VIL\_HPV45 STANDARD; PRT; 539 AA.  
 ID VIL\_HPV45  
 AC P36741;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Major capsid protein LI.  
 GN LI.  
 OS Human papillomavirus type 45.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 CC Papillomavirus.  
 OX NCBI\_TaxID=10593;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94265501; PubMed=8205838;  
 RA Delius H., Hofmann B.;  
 RT "Primer-directed sequencing of human papillomavirus types";  
 RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
 CC "This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))."  
 CC EMBL; X74479; CAAS2578.1; -  
 DR PIR; S36566; S36566.

DR InterPro: IPR002210; PV\_capsid\_L1.  
 DR Pfam: PF005500; Late\_protein\_L1; 1.  
 DR PRINTS: PR00865; HPVcapsidL1.  
 DR Prodom: PD000544; PV\_capsid\_L1; 1.  
 KW Coat protein; late protein.  
 SQ SEQUENCE 539 AA; 60310 MW; 9D79F6B804234E5E CRC64;  
 Query Match 89.6%; Score 2450; DB 1; Length 539;  
 Best Local Similarity 87.3%; Pred. No. 1.1e-204;  
 Matches 448; Conservative 30; Mismatches 29; Indels 6; Gaps 3;

QY 1 MALMRPSNTYLLPPPSVARYVNDYTRKSIYFHAGSSRLTYGNFYFRYPAGGK 59  
 |||||||  
 DB 27 MALMRPSNTYLLPPPSVARYVNDYTRKSIYFHAGSSRLTYGNFYFRYPAGGK 86  
 |||||||  
 QY 60 DIPKVSAYQYRVFVQVLPDPNKKFGLPDNSIYNPEQRLVWACAGVEIGROPGLGSLG 119  
 |||||||  
 DB 87 QAVRVSAYQYRVFVQVLPDPNKKFGLPDNSIYNPEQRLVWACAGVEIGROPGLGSLG 146  
 |||||||  
 QY 120 HPFYKLLDTESSHAATSNVSEVDVNDVSDYKQVQLCICGAPALGEHMAKGTACKSRP 179  
 |||||||  
 DB 147 HPFYKLLDTESSHAATSNVSEVDVNDVSDYKQVQLCICGAPALGEHMAKGTACKSRP 206  
 |||||||  
 QY 180 LSOGDCPRLKNTYLEDDGMVDYGYGAMDESTLDQTKCEVPLDQICQICKYPPYLQMSA 239  
 |||||||  
 DB 207 LSOGDCPRLKNTYLEDDGMVDYGYGAMDESTLDQTKCEVPLDQICQICKYPPYLQMSA 266  
 |||||||  
 QY 240 PYGDSMFECRLRQDLFAHFNNRAGTMDYVPOSILYIKGT - GMASPGSCYSPSPSG 297  
 |||||||  
 DB 267 PYGDSMFECRLRQDLFAHFNNRAGTMDYVPOSILYIKGT - GMASPGSCYSPSPSG 326  
 |||||||  
 QY 298 STVSDSOLFENKPYWHLKAGCHNNGICWHNOLFVTVVDTTRSTNLITCASTQSPVPGQYD 357  
 |||||||  
 DB 327 STVSDSOLFENKPYWHLKAGCHNNGICWHNOLFVTVVDTTRSTNLITCASTQSPVPGQYD 386  
 |||||||  
 QY 358 ATKFKQSRHVEEYDLOFIFOLCTITLADYMSYIHSNSSLIEDMNGVPPPTSLVDY 417  
 |||||||  
 DB 387 ATKFKQSRHVEEYDLOFIFOLCTITLADYMSYIHSNSSLIEDMNGVPPPTSLVDY 446  
 |||||||  
 QY 418 TYRFVQSVATTCQKDAAPAEKNDYDKLKFNNVDLKEKESLDLDQYPLGRKFLVQAGLR 477  
 |||||||  
 DB 447 TYRFVQSVATTCQKDAAPAEKNDYDKLKFNNVDLKEKESLDLDQYPLGRKFLVQAGLR 506  
 |||||||  
 QY 478 KPTIGPKRRKAP--ATTSSKPAKRVVRARK 507  
 |||||||  
 DB 507 KPTIGPKRRKAP--ATTSSKPAKRVVRARK 539  
 |||||||

RESULT 3  
 VIL\_HPV70 STANDARD; PRT; 504 AA.  
 ID VIL\_HPV70  
 AC P50793;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Major capsid protein LI.  
 GN LI.  
 OS Human papillomavirus type 70.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 CC Papillomavirus.  
 OX NCBI\_TaxID=39457;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96249586; PubMed=8815087;  
 RA Forslund O., Hansson B.G.;  
 RT "Human papillomavirus type 70 genome cloned from overlapping PCR products: complete nucleotide sequence and genomic organization";  
 RL J. Clin. Microbiol. 34:802-809(1996).  
 CC "This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))."  
 CC EMBL; X74479; CAAS2578.1; -  
 DR PIR; S36566; S36566.



```
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 39.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_Taxid=10588;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9135017; PubMed=1847266;
RA Volpers C., Streck R.E.;
RT "Genome organization and nucleotide sequence of human papillomavirus
type 39."
RL Virology 181:419-423(1991).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M62849; AAA47056.1; -
DR PIR; H38502; P1W39.
DR InterPro; IPR002210; PV_capsid_L1.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KW Coat protein; late protein.
SQ
SEQUENCE 505 AA; 56604 MW; 73854P9F1C56BCD CRC64;

Query Match 78.7%; Score 2151; DB 1; Length 505;
Best Local Similarity 77.3%; Pred. No. 8,6e-179;
Matches 391; Conservative 48; Mismatches 61; Indels 6; Gaps 3;

QY 1 MALTSPDNTVYLPPEPVAVRYNTDVTYRTSIFAHAGSSRLTVGNPFYRVPAAGGKQ 60
DB 1 MAMRRSDSMVYLPPEPVAVRYNTDVTYRTSIFAHAGSSRLTVGNPFYRVPAAGGKQ 60
QY 61 DIPVSAIYQYRVFVYLPDPNKKFGLPDNSIYNPETORLYMACGVEIGRGQPLAVGLSGH 120
DB 61 DIPVSAIYQYRVFVYLPDPNKKFGLPDNSIYNPETORLYMACGVEIGRGQPLAVGLSGH 120
QY 121 PFYNNKLDTESSHAATSNVSEVDVNVSVYKOTQLCICAPAIGEHMAKGTACKSRPL 180
DB 121 PFYNNKLDTESSHAATSNVSEVDVNVSVYKOTQLCICAPAIGEHMAKGTACKSRPL 180
QY 121 PLYNRQDTESSHAATSNVSEVDVNVSVYKOTQLCICAPAIGEHMAKGTACKSRPL 179
DB 121 PLYNRQDTESSHAATSNVSEVDVNVSVYKOTQLCICAPAIGEHMAKGTACKSRPL 179
QY 181 SOGCPPELEKNTVLEGGDMDVDTGYGAMDSITLDDTKCEVPLDTCOSICKPYLQMSAD 240
DB 181 SOGCPPELEKNTVLEGGDMDVDTGYGAMDSITLDDTKCEVPLDTCOSICKPYLQMSAD 240
QY 180 STGCAPPELEKNTVLEGGDMDVDTGYGAMDSITLDDTKCEVPLDTCOSICKPYLQMSAD 239
DB 180 STGCAPPELEKNTVLEGGDMDVDTGYGAMDSITLDDTKCEVPLDTCOSICKPYLQMSAD 239
QY 241 PYGDSMFCLRRDLFARHFNWNRAGTMGDVYPOSILYKGTGMASPGSCYSPSPSGSIY 300
DB 241 PYGDSMFCLRRDLFARHFNWNRAGTMGDVYPOSILYKGTGMASPGSCYSPSPSGSIY 300
QY 240 VYGDSEMFCLRRDLFARHFNWNRAGTMGDVYPOSILYKGTGMASPGSCYSPSPSGSIY 299
DB 240 VYGDSEMFCLRRDLFARHFNWNRAGTMGDVYPOSILYKGTGMASPGSCYSPSPSGSIY 299
QY 301 TSDSOLFENKPYWLKKAAGHNGICMHNQLEFVTVVDTSTRNLITCASTQSPVPGQYDATK 360
DB 301 TSDSOLFENKPYWLKKAAGHNGICMHNQLEFVTVVDTSTRNLITCASTQSPVPGQYDATK 360
QY 300 TSDSOLFENKPYWLKKAAGHNGICMHNQLEFVTVVDTSTRNLITCASTQSPVPGQYDATK 359
DB 300 TSDSOLFENKPYWLKKAAGHNGICMHNQLEFVTVVDTSTRNLITCASTQSPVPGQYDATK 359
QY 361 FKQYSRHVEEYDLOFIFOLCTITLTADVMYSIHSMNSIIEDMNFVGPVPPPTSLVDTYR 420
DB 361 FKQYSRHVEEYDLOFIFOLCTITLTADVMYSIHSMNSIIEDMNFVGPVPPPTSLVDTYR 420
QY 360 FKEVTRHVEEYDLOFIFOLCTITLTADVMYSIHSMNSIIEDMNFVGPVPPPTSLVDTYR 419
DB 360 FKEVTRHVEEYDLOFIFOLCTITLTADVMYSIHSMNSIIEDMNFVGPVPPPTSLVDTYR 419
QY 421 FVQSVATTCQKDAAPAKNKPYPYDLKTFMNVDLKEKFSLDLQYPLGKFLYQAGLRKRP 480
DB 421 FVQSVATTCQKDAAPAKNKPYPYDLKTFMNVDLKEKFSLDLQYPLGKFLYQAGLRKRP 480
QY 481 IGPKRSPASATSSKPA---KRVAV 503
DB 481 IGPKRSPASATSSKPA---KRVAV 503
QY 480 IGPKRSPASATSSKPA---KRVAV 503
DB 480 IGPKRSPASATSSKPA---KRVAV 503
```

```
RESULT 6
VLI_HPV68 STANDARD; PRT; 505 AA.
AC P54669;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 68.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_Taxid=45240;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97060129; PubMed=8904450;
RA Longuet M., Beaudenon S., Orth G.;
RT "Two novel genital human papillomavirus (HPV) types, HPV68 and HPV70,
RL J. Clin. Microbiol. 34:738-744(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X67161; CAA47634.1; -
DR InterPro; IPR002210; PV_capsid_L1.
DR Pfam; PF00500; late_protein_L1; 1.
DR PRINTS; PR00865; HPVcapsidL1.
DR ProDom; PD000544; PV_capsid_L1; 1.
KW Coat protein; late protein.
SQ
SEQUENCE 505 AA; 56693 MW; 246CDB4DAB2A1997 CRC64;

Query Match 78.3%; Score 2141; DB 1; Length 505;
Best Local Similarity 75.9%; Pred. No. 6,4e-178;
Matches 382; Conservative 56; Mismatches 65; Indels 0; Gaps 0;

QY 1 MALTSPDNTVYLPPEPVAVRYNTDVTYRTSIFAHAGSSRLTVGNPFYRVPAAGGKQ 60
DB 1 MAMRRSDSMVYLPPEPVAVRYNTDVTYRTSIFAHAGSSRLTVGNPFYRVPAAGGKQ 60
QY 61 DIPVSAIYQYRVFVYLPDPNKKFGLPDNSIYNPETORLYMACGVEIGRGQPLAVGLSGH 120
DB 61 DIPVSAIYQYRVFVYLPDPNKKFGLPDNSIYNPETORLYMACGVEIGRGQPLAVGLSGH 120
QY 121 PFYNNKLDTESSHAATSNVSEVDVNVSVYKOTQLCICAPAIGEHMAKGTACKSRPL 180
DB 121 PFYNNKLDTESSHAATSNVSEVDVNVSVYKOTQLCICAPAIGEHMAKGTACKSRPL 180
QY 121 PLYNRQDTESSHAATSNVSEVDVNVSVYKOTQLCICAPAIGEHMAKGTACKSRPL 180
DB 121 PLYNRQDTESSHAATSNVSEVDVNVSVYKOTQLCICAPAIGEHMAKGTACKSRPL 180
QY 181 SOGCPPELEKNTVLEGGDMDVDTGYGAMDSITLDDTKCEVPLDTCOSICKPYLQMSAD 240
DB 181 SOGCPPELEKNTVLEGGDMDVDTGYGAMDSITLDDTKCEVPLDTCOSICKPYLQMSAD 240
QY 180 STGCAPPELEKNTVLEGGDMDVDTGYGAMDSITLDDTKCEVPLDTCOSICKPYLQMSAD 239
DB 180 STGCAPPELEKNTVLEGGDMDVDTGYGAMDSITLDDTKCEVPLDTCOSICKPYLQMSAD 239
QY 241 PYGDSMFCLRRDLFARHFNWNRAGTMGDVYPOSILYKGTGMASPGSCYSPSPSGSIY 300
DB 241 PYGDSMFCLRRDLFARHFNWNRAGTMGDVYPOSILYKGTGMASPGSCYSPSPSGSIY 300
QY 240 VYGDSEMFCLRRDLFARHFNWNRAGTMGDVYPOSILYKGTGMASPGSCYSPSPSGSIY 299
DB 240 VYGDSEMFCLRRDLFARHFNWNRAGTMGDVYPOSILYKGTGMASPGSCYSPSPSGSIY 299
QY 301 TSDSOLFENKPYWLKKAAGHNGICMHNQLEFVTVVDTSTRNLITCASTQSPVPGQYDATK 360
DB 301 TSDSOLFENKPYWLKKAAGHNGICMHNQLEFVTVVDTSTRNLITCASTQSPVPGQYDATK 360
QY 300 TSDSOLFENKPYWLKKAAGHNGICMHNQLEFVTVVDTSTRNLITCASTQSPVPGQYDATK 359
DB 300 TSDSOLFENKPYWLKKAAGHNGICMHNQLEFVTVVDTSTRNLITCASTQSPVPGQYDATK 359
QY 361 FKQYSRHVEEYDLOFIFOLCTITLTADVMYSIHSMNSIIEDMNFVGPVPPPTSLVDTYR 420
DB 361 FKQYSRHVEEYDLOFIFOLCTITLTADVMYSIHSMNSIIEDMNFVGPVPPPTSLVDTYR 420
QY 360 FKEVTRHVEEYDLOFIFOLCTITLTADVMYSIHSMNSIIEDMNFVGPVPPPTSLVDTYR 419
DB 360 FKEVTRHVEEYDLOFIFOLCTITLTADVMYSIHSMNSIIEDMNFVGPVPPPTSLVDTYR 419
QY 421 FVQSVATTCQKDAAPAKNKPYPYDLKTFMNVDLKEKFSLDLQYPLGKFLYQAGLRKRP 480
DB 421 FVQSVATTCQKDAAPAKNKPYPYDLKTFMNVDLKEKFSLDLQYPLGKFLYQAGLRKRP 480
```



DB 421 YLQSAITCQKAPAPVKKDPYDGLNFVNVDLKEKFSSELDQPLGRKFLDQAGVRRPT 480  
QY 481 IGRKRSAPSATSSSKPARKRV 503  
DB 481 IGRKRSATTTSTSKHRRKV 503

RESULT 7  
V1L\_HPV29  
ID V1L\_HPV29 STANDARD: PRT: 503 AA.  
AC P50792:  
DB 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DE 01-OCT-1996 (Rel. 34, Last annotation update)  
Major capsid protein L1.  
GN L1.  
OS Human papillomavirus type 29.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxId=37112;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Delius H.;  
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 316-467 FROM N.A.  
RX MEDLINE-95052821; PubMed-7963696;  
RA Bernard H.U., Chan S.Y., Manos M.M., Ong C.R., Villa L.L.,  
RA Delius H., Peyton C.L., Bauer H.M., Wheeler C.M.,  
RT "Identification and assessment of known and novel human  
RT papillomaviruses by polymerase chain reaction amplification,  
RT restriction fragment length polymorphisms, nucleotide sequence, and  
RT phylogenetic algorithms.";  
RL J. Infect. Dis. 170:1077-1085(1994).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U31784; AAA79435.1; -  
DR EMBL: U12503; AAA67247.1; -  
DR InterPro: IPR002210; PV\_capsid.L1.  
DR Pfam: PF00500; late\_protein.L1.  
DR PRINTS: PR00865; HPVcapsid.L1.  
DR ProDom: PD000544; PV\_capsid.L1.  
DR Coar protein; late protein.  
SQ SEQUENCE 503 AA; 56281 MW; 8D2E36DAABD30847 CRC64;

Query Match 71.6%; Score 1956; DB 1; Length 503;  
Best Local Similarity 69.3%; Pred. No. 7, 1e-162;  
Matches 348; Conservative 70; Mismatches 82; Indels 2; Gaps 1;

QY 1 MALMRPSDNTVYLPPPSVAVVNTDDYVTRTSIFYHAGSSRLITGVGNPFYRYPAGGNGK 60  
DB 1 MALMRSSDNLVYLPPTPVSKVISTDDYVTRTNIYYAGSSRLITGVHPIYSIPKSGNKV 60  
QY 61 DIPKSAOYRYFRVQLPDPNKFGLPDNSITYPEQRLVWACAGVEIGGQPLGVLGSH 120  
DB 61 DIPKSAOYRYFRVQLPDPNKFGLPDAIYNPBEARLWMACTGVEVGQPLGVLGSH 120  
QY 121 PFYRKLDPTSSHAATSNVSEVDVNVSDYKOTQCTICGAPATGEHNAKGTACKSRPL 180  
DB 121 PFYRKLDPTSSHAATSNVSEVDVNVSDYKOTQCTICGAPATGEHNAKGTACKSRPL 180  
QY 121 PFYRKLDPTSSHAATSNVSEVDVNVSDYKOTQCTICGAPATGEHNAKGTACKSRPL 180  
DB 121 PFYRKLDPTSSHAATSNVSEVDVNVSDYKOTQCTICGAPATGEHNAKGTACKSRPL 180  
QY 181 SOGDPEPELKTIVLEDGMDVDTGYGAMDSTLQDTKCEVPLDICSICKYFDYLOMSAD 240  
DB 181 AAGDCPPLELMTTHIEDGMDVDTGYGAMDFAALQVKKSDVPLDICSTCKYFDYLOMSAD 240  
QY 241 PYGDSMFELRREQLFARHFNRAGTMDGTFQSLYIKGTGMRASPGSCVYSPSGSIV 300

DB 241 PYGDSMFELRREQLFARHFNRAGTMDGTFQSLYIKGTGMRASPGSIV 300  
QY 301 TSDSOLFENRPMYLRKQGNNGCWANQLFVYVDTBSTNTLTICASTOSPVPQYDATK 360  
DB 301 TSDSOLFENRPMYLRKQGNNGCWANQLFVYVDTBSTNTLTICASTOSPVPQYDATK 360  
QY 361 FKQYSRVVEYDQIFQCTITLIFADVNSYKSHNSSTLIEDMNFQPPPTSLVDYR 420  
DB 361 IKEYLRHGEFYDQIFQCTITLIFADVNSYKSHNSSTLIEDMNFQPPPTSLVDYR 420  
QY 421 FVQSAITCQKAPAPVKKDPYDGLNFVNVDLKEKFSSELDQPLGRKFLDQAGVRRPT 480  
DB 421 FVQSAITCQKAPAPVKKDPYDGLNFVNVDLKEKFSSELDQPLGRKFLDQAGVRRPT 480  
QY 481 IGRKRSAPSATSSSKPARKRV 502  
DB 481 IGRKRSAPSATSSSKPARKRV 502

RESULT 8  
V1L\_HPV03  
ID V1L\_HPV03 STANDARD: PRT: 532 AA.  
AC P36731;  
DB 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DE 01-OCT-1996 (Rel. 34, Last annotation update)  
Major capsid protein L1.  
GN L1.  
OS Human papillomavirus type 3.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxId=10614;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-94265501; PubMed-8205838;  
RA Delius H.;  
RT "Primer-directed sequencing of human papillomavirus types.";  
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-29 IS THE INITIATOR.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X74462; CA52474.1; -  
DR PIR: S36554; S36554.  
DR InterPro: IPR002210; PV\_capsid.L1.  
DR Pfam: PF00500; late\_protein.L1.  
DR PRINTS: PR00865; HPVcapsid.L1.  
DR ProDom: PD000544; PV\_capsid.L1.  
DR Coar protein; late protein.  
SQ SEQUENCE 532 AA; 59194 MW; A0AFB87DBEE98B27 CRC64;

Query Match 70.2%; Score 1917.5; DB 1; Length 532;  
Best Local Similarity 68.8%; Pred. No. 1, 7e-158;  
Matches 351; Conservative 65; Mismatches 85; Indels 9; Gaps 4;

QY 1 MALMRPSDNTVYLPPPSVAVVNTDDYVTRTSIFYHAGSSRLITGVGNPFYRYPAGGNGK 60  
DB 1 MALMRSSDNLVYLPPTPVSKVISTDDYVTRTNIYYAGSSRLITGVHPIYSIPKSGNKV 60  
QY 29 MALMRSSDNLVYLPPTPVSKVISTDDYVTRTNIYYAGSSRLITGVHPIYSIPKSGNKV 88  
DB 29 MALMRSSDNLVYLPPTPVSKVISTDDYVTRTNIYYAGSSRLITGVHPIYSIPKSGNKV 88  
QY 61 DIPKSAOYRYFRVQLPDPNKFGLPDNSITYPEQRLVWACAGVEIGGQPLGVLGSH 120  
DB 61 DIPKSAOYRYFRVQLPDPNKFGLPDAIYNPBEARLWMACTGVEVGQPLGVLGSH 148  
QY 121 PFYRKLDPTSSHAATSNVSEVDVNVSDYKOTQCTICGAPATGEHNAKGTACKSRPL 180  
DB 121 PFYRKLDPTSSHAATSNVSEVDVNVSDYKOTQCTICGAPATGEHNAKGTACKSRPL 180  
QY 149 PFYRKLDPTSSHAATSNVSEVDVNVSDYKOTQCTICGAPATGEHNAKGTACKSRPL 207

QY	181	SQDCPPELEKNTVLEDCGMVDTGCADEFTLDQTEKVEPLDLCOSICKPVDLKQMSD	240
Db	208	SPEDCEPPELILNAPLIDGSDMVDYTGCADEFGILQSKNSDYDELIDQITCKRPDLIGMAE	267
QY	241	PYDSMFECRLRRQOLFARHFNNRAGTMGDYVPQSLYIKG--TGMRAFPSGCVTSPSPSG	297
Db	268	PYDSMFELRLKEQLFARHFLNRAMADDPDALYIGDSSQSGRKIGSAVYCCPRPSSG	327
QY	298	SIYTSQSOLFENKPYWLHKAQGHNNICWNHOLEFVVVDYTRSTNLTICASTQSPVPOYD	357
Db	328	SWMTSTPOLFENKPYWLRRAGGHNNICWANOGLFVVVDYTRSTNMTLCVSHET--SMTYD	385
QY	358	AATFCKRYSHRYEYDQFTEQLCTTLTADWASYIHSNNSLLEDMANGVPRPTTSLYD	417
Db	386	AATFCKYLRHGEYDQFTEQLCKYTLPELIMAYLHTNNSLLEDMWNGTLPLPSTSLD	445
QY	418	TYRFVSVAITCQDAAPAEKNDPVDKLTFWNVNDLEKEFSLDDQYPLGRKFLYQAOLRR	477
Db	446	TYRFLNLSAITCQKDAPEYKQDPYAKLNFVNDLKDRLFSLDSQPLGRKFLMQLGVT	505
QY	478	KPTIGPRKRSAPATSTSSKPAKRYVRARK	507
Db	506	RSSISVRRK--SATTSTTAALKRRRK	532

RESULT 9		
VL1_HPV28	STANDARD:	PRI: 502 AA.
ID	VL1_HPV28	
AC	P30791;	
DT	01-OCT-1996 (Rel. 34, Created)	
DT	01-OCT-1996 (Rel. 34, Last sequence update)	
DT	01-OCT-1996 (Rel. 34, Last annotation update)	
DE	Major capsid protein L1.	
GN	L1.	
OS	Human papillomavirus type 28.	
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;	
OC	Papillomavirus.	
OX	NCBI_TaxID=37111;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Dellius H.;	
RT	Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.	
RL	[2]	
RP	SEQUENCE OF 316-464 FROM N.A.	
RX	MEDLINE=95052821; Pubmed=796396;	
RA	Bernard H.U., Chan S.Y., Manos M.M., Ong C.K., Villa L.L.,	
RA	Dellius H., Peyton C.L., Bauer H.M., Wheeler C.M.;	
RT	"Identification and assessment of known and novel human	
RT	papillomaviruses by polymerase chain reaction amplification,	
RT	restriction fragment length polymorphisms, nucleotide sequence, and	
RT	phylogenetic algorithms."	
RL	J. Infect. Dis. 170:1077-1085(1994).	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/sib-sib.ch">http://www.isb-sib.ch/announce/sib-sib.ch</a> ).	
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).	
CC	-----	
DR	EMBL: U31783; AAA79428.1; -.	
DR	EMBL: U12502; AAA67246.1; -.	
DR	InterPro: IPR002210; PY_capsid_L1.	
DR	Pfam: PF00500; Late_protein_L1; 1.	
DR	PRINTS: PR00865; HPVcapsIDL1.	
DR	ProDom: PD000544; PY_capsid_L1; 1.	
KW	Coat protein; Late protein.	
SEQUENCE	502 AA; 55902 MW; 38D6AD812AF310A CRC64;	
Query Match	70.0%; Score 1912.5; DB 1; Length 502;	
Best Local Similarity	68.3%; Pred. NO. 4.2e-158;	

[illegible]

```

RESULT 10
VRL_HPV10
ID VRL_HPV10 STANDARD; PRT; 531 AA.
AC P36732;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Major capsid protein L1.
GN L1.
OS Human papillomavirus type 10.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_Taxid=10603;
[1]
RN RP SEQUENCE FROM N.A.
RX MEDLINE=9426501; Pubmed=8205838;
RA Delius H., Hofmann B.;
RT "Primer-directed sequencing of human papillomavirus types.*";
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC -!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-29 IS THE INITIATOR.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X74465; CAAS2494.1; -.
DR PIR: S36537; S36537.
DR InterPro: IPR002210; PY_capsid_L1.
DR Pfam: PF00500; late_protein_L1.
DR PRINTS: PR00865; HPVcapsidL1.
DR

```

DR Prodom: PD000544; PV\_capsid\_L1; 1.

KW Coat protein; Late protein.  
SEQUENCE 531 AA; 59011 MW; EDC9BADCA7A67F6 CRC64;

Query Match 69.8%; Score 1909; DB 1; Length 531;  
Best Local Similarity 67.9%; Pred. No. 9.2e-158;  
Matches 342; Conservative 70; Mismatches 90; Indels 2; Gaps 2;

QY 1 MALMRSDNTVYLPPPSVARVNTDDYVTRTSIFYHAGSSRLTLTGNGPYFRVAGGNKQ 60  
DB 29 MALMRSSDNLVYLPPPSVKVLTDDYVTRNTIYYAGSRLLTGVHPFPIPKSSNNKY 88  
QY 61 DIPKVSAYQYRVFVQLPDPNKGFLPDNSIYNPETORLWACAGVEIGRQPLGVLSGH 120  
DB 89 DVPKVSAYQYRVFVRLPDPNKGFLPDARVYNDARLWACAGVEIGRQPLGVLSGH 148  
QY 121 PFYNKLDDESSHAATSNVSEEDVADNVSDYKOTQCLICGAPALGEHNAKGTACKSRPL 180  
DB 149 PLTNKLEDEENSIAGPIGQDSRDNIYDNKOTQCLICGTPPEHNGKGTPCRNP- 207  
QY 181 SQGDCPPELEKNTVLEDDGVDTGYGAMDFSLQDTKCEVPLDQCOSICKYPPYLQMSAD 240  
DB 208 AOGDCPPELEITSPIODGDVDTGYGAMPETALQLNKSDVPIDICOSTCKYPPYLQMAE 267  
QY 241 PYGDSMEFCLRRQQLFARHFNAGTMDTPVOSLYIKGTGMARSGSCVSPSGSIY 300  
DB 268 DVPKVSAYQYRVFVRLPDPNKGFLPDARVYNDARLWACAGVEIGRQPLGVLSGH 327  
QY 301 TSDSOLFENKPYMLHKAQGNHNGICWGNOLFYVVDTRSTNLTCASTOSPPVGOYDATK 360  
DB 328 TSEAQLFENKPYMLHKAQGNHNGICWGNOLFYVVDTRSTNLTCASTOSPPVGOYDATK 387  
QY 361 FKOYSHVEEYDQFIFQCTITLTDVNSYIHSNSSLIEDMNFVPPPTSLVDYR 420  
DB 388 FKEYLHGEYDQFIFQCTITLTDVNSYIHSNSSLIEDMNFVPPPTSLVDYR 447  
QY 421 FVQSVAITCOKDAAPAKENDPYDKLKFVNVDLKEKFSLDLDQYPLGRKFLVQGLRRKP 480  
DB 448 FVSSSLAITCOKDPPPEKEDPYAKLNFVNDLKDPSLDLSOPPLGRKFLVQGLRRSA 507  
QY 481 IGPKRSAPSATTSKPAKRVYR 504  
DB 508 VSVRKRPAATSA-TGSFAAKRRRK 530

RESULT 11  
VIL\_HPV30 STANDARD; PRT; 508 AA.  
ID VIL\_HPV30  
AC 002515;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Major capsid protein L1.  
GN L1.  
OS Human papillomavirus type 30.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_Taxid=10611;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94265501; Pubmed=8205838;  
RA Dellus H., Hofmann B.;  
RT "Primer-directed sequencing of human papillomavirus types.";  
RN Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
RN (2)  
RP SEQUENCE OF 309-352 FROM N.A.  
RX MEDLINE=92407963; Pubmed=1326639;  
RA Chan S.Y., Bernard H.U., Ong C.K., Chan S.P., Birgit H., Dellus H.;  
RT "Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and  
variants: a showcase for the molecular evolution of DNA viruses.";  
RN J. Virol. 66:5714-5725(1992).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: X74474; CA52548.1; -;  
DR EMBL: M96279; AAA47018.1; -;  
DR PIR: S36508; S36508.  
DR InterPro: IPR002210; PV\_capsid\_L1.  
DR Pfam: PF00500; Late-protein\_L1.  
DR PRINTS: PR00865; HPVcapsidL1.  
DR Prodom: PD000544; PV\_capsid\_L1; 1.  
KW Coat protein; Late protein.  
SQ SEQUENCE 508 AA; 56667 MW; 62F359F257714748 CRC64;

Query Match 68.8%; Score 1879; DB 1; Length 508;  
Best Local Similarity 67.1%; Pred. No. 3.5e-155;  
Matches 337; Conservative 82; Mismatches 81; Indels 2; Gaps 1;

QY 1 MALMRSDNTVYLPPPSVARVNTDDYVTRTSIFYHAGSSRLTLTGNGPYFRVAGGNKQ 60  
DB 8 MALMRSETKVYLPPPSKVPTDLYVKTNTIFYHAGSSRLTAVGHPYSSISKAGNSKT 67  
QY 61 DIPKVSAYQYRVFVQLPDPNKGFLPDNSIYNPETORLWACAGVEIGRQPLGVLSGH 120  
DB 68 DVPKVSAYQYRVFVRLPDPNKGFLPDVNFNEGRVRLWACAGVEIGRQPLGVLSGH 127  
QY 121 PFYNKLDDESSHAATSNVSEEDVADNVSDYKOTQCLICGAPALGEHNAKGTACKSRPL 180  
DB 128 PFENKLDDESSSTIANQDAEDSRDNIYDNKOTQCLICGTPPEHNGKGTACKSRAP 187  
QY 181 SQGDCPPELEKNTVLEDDGVDTGYGAMDFSLQDTKCEVPLDQCOSICKYPPYLQMSAD 240  
DB 188 AOGDCPPELELVNSPIDGDVDTGYGAMDFKTLQESKSDVPLDISOSTCKYPPYLQMSAD 247  
QY 241 PYGDSMEFCLRRQQLFARHFNAGTMDTPVOSLYIKGTGMARSGSCVSPSGSIY 300  
DB 248 AYGDMSWVFLRRQQLFARHFNAGTMDTPVOSLYIKGTGMARSGSCVSPSGSIY 307  
QY 301 TSDSOLFENKPYMLHKAQGNHNGICWGNOLFYVVDTRSTNLTCASTOSPPVGOYDATK 360  
DB 308 TSEAQLFENKPYMLHKAQGNHNGICWGNOLFYVVDTRSTNLTCASTOSPPVGOYDATK 365  
QY 361 FKOYSHVEEYDQFIFQCTITLTDVNSYIHSNSSLIEDMNFVPPPTSLVDYR 420  
DB 366 IKQYVHVEEYDQFIFQCTITLTDVNSYIHSNSSLIEDMNFVPPPTSLVDYR 425  
QY 421 FVQSVAITCOKDAAPAKENDPYDKLKFVNVDLKEKFSLDLDQYPLGRKFLVQGLRRKP 480  
DB 426 YVKSIAITCOKDPPPEKEDPYAKLNFVNDLKDPSLDLSOPPLGRKFLVQGLRRKTS 485  
QY 481 IGPKRSAPSATTSKPAKRVYR 502  
DB 486 TTKRKSAPSSSTSPSAKRRK 507

RESULT 12  
VIL\_HPV26 STANDARD; PRT; 503 AA.  
ID VIL\_HPV26  
AC P36735;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Major capsid protein L1.  
GN L1.  
OS Human papillomavirus type 26.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_Taxid=31549;  
RN (1)  
RP SEQUENCE FROM N.A.

RA MEDLINE=94265501; PubMed=8205838;  
RA Delius H., Hofmann B.;  
RT "Primer-directed sequencing of human papillomavirus types.";  
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; X74472; CAA5235.1; -  
DR PIR; S36549; S36549.  
DR InterPro; IPR002210; PV\_capsid\_L1.  
DR Pfam; PF00500; late\_protein\_L1; 1.  
DR PRINTS; PR00865; HPVcapsidL1.  
DR ProDom; PD000544; PV\_capsid\_L1; 1.  
DR Coa protein; Late protein.  
SQ SEQUENCE 503 AA; 56328 MW; A0706D12F425BE80 CRC64;  
  
Query Match 68.1%; Score 1861.5; DB 1; Length 503;  
Best Local Similarity 67.3%; Pred. No. 1,le-153;  
Matches 337; Conservative 70; Mismatches 91; Indels 3; Gaps 3;  
  
QY 1 MALMRPSDNVYLLPPPPSVARYVNTDVTYRTSIFVHAGSSRLTLVGNPFYRVPAGGKQ 60  
DB 1 MALMRPSDNVYLLPPPPSVARYVNTDVTYRTSIFVHAGSSRLTLVGNPFYRVPAGGKQ 59  
QY 61 DIPVSAQYRFRVQLPDPNKFGLPDNSIYNPEFQRLVWACAGVEIGRGPPLVGLSGH 120  
DB 60 EIPVSAQYRFRVQLPDPNKFGLPDNSIYNPEFQRLVWACAGVEIGRGPPLVGLSGH 119  
QY 121 PFYNNKLDITDESSHAATSNVEDVNDVSVYKQOTOLCIGCAPAIGEHMAGTACKSRPL 180  
DB 120 PLFNKLDITDESSHAATSNVEDVNDVSVYKQOTOLCIGCAPAIGEHMAGTACKSRPL 179  
QY 181 SQGDCCPLELKNVLEDEGMDVDTGYGAMDEFTLDTCKEVPDLTQSCICKYPDYLOMSAD 240  
DB 180 SQGDCCPLELKNVLEDEGMDVDTGYGAMDEFTLDTCKEVPDLTQSCICKYPDYLOMSAD 239  
QY 241 PYGDSMFCLREQLFARHFNRACTMGDTVPQSLYIKGT-GMARSPSCVYSPSPSGS 299  
DB 240 TYGDSMFCLREQLFARHFNRACTMGDTVPQSLYIKGT-GMARSPSCVYSPSPSGS 299  
QY 300 VTSQSOLFENKPYWLHKAQGHNNNGICWNLFTVVDTRSTNLTICASTQSPVPGQYDAT 359  
DB 300 VTSQSOLFENKPYWLHKAQGHNNNGICWNLFTVVDTRSTNLTICASTQSPVPGQYDAT 359  
QY 360 KFKQYSRHVEYDIQIFQLCTITLADVMSYIHSNNSSTLEDMNFGVPPPTSLVDY 419  
DB 360 DYKQFIRHVEYDIQIFQLCTITLADVMSYIHSNNSSTLEDMNFGVPPPTSLVDY 419  
QY 420 RFVQSVATTCOKDAAPAKNDPYDKLFWNDLKEKESLDDOYPLGKFLVQAGLRKP 479  
DB 420 RFVQSVATTCOKDAAPAKNDPYDKLFWNDLKEKESLDDOYPLGKFLVQAGLRKP 479  
QY 480 TIGPRKRSAPSATTSKPAKR 500  
DB 480 TIGPRKRSAPSATTSKPAKR 499  
DB 480 KLG-TKRLPSTSSSTTKRKR 499  
  
RESULT 13  
V1 HPV61 STANDARD; PRT; 505 AA.  
AC P50822; 080954;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Major capsid protein L1.  
GN L1.  
OS Human papillomavirus type 61.

OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OC NCBI\_TaxID=37116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Delius H.;  
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 317-468 FROM N.A.  
RX MEDLINE=95052821; PubMed=7963696;  
RA Bernard H.U., Chan S.Y., Manos M.M., Ong C.K., Villa L.L.,  
RT "Identification and assessment of known and novel human  
RT papillomaviruses by polymerase chain reaction amplification,  
RT restriction fragment length polymorphisms, nucleotide sequence, and  
RT phylogenetic algorithms.";  
RL Infect. Dis. 170:1077-1085(1994).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; U31793; AAA79498.1; -  
DR EMBL; U12500; AAA67244.1; -  
DR InterPro; IPR002210; PV\_capsid\_L1.  
DR Pfam; PF00500; late\_protein\_L1; 1.  
DR PRINTS; PR00865; HPVcapsidL1.  
DR ProDom; PD000544; PV\_capsid\_L1; 1.  
DR Coa protein; Late protein.  
FT CONFLICT 374 374 0 -> P (IN REF. 2).  
FT CONFLICT 430 430 0 -> L (IN REF. 2).  
SQ SEQUENCE 505 AA; 56710 MW; EB56C4A60E94C853 CRC64;  
  
Query Match 67.9%; Score 1855.5; DB 1; Length 505;  
Best Local Similarity 66.5%; Pred. No. 3,8e-153;  
Matches 337; Conservative 69; Mismatches 94; Indels 7; Gaps 5;  
  
QY 1 MALMRPSDNVYLLPPPPSVARYVNTDVTYRTSIFVHAGSSRLTLVGNPFYRVPAGG--GN 58  
DB 1 MALMRPSDNVYLLPPPPSVARYVNTDVTYRTSIFVHAGSSRLTLVGNPFYRVPAGG 60  
QY 59 KODIPKYSAYQYRFRVQLPDPNKFGLPDNSIYNPEFQRLVWACAGVEIGRGPPLVGLSGH 118  
DB 61 KNTIPKYSAYQYRFRVQLPDPNKFGLPDNSIYNPEFQRLVWACAGVEIGRGPPLVGLSGH 120  
QY 119 GHPIYNNKLDITDESSHAATSNVEDVNDVSVYKQOTOLCIGCAPAIGEHMAGTACKSR 178  
DB 121 GHPIYNNKLDITDESSHAATSNVEDVNDVSVYKQOTOLCIGCAPAIGEHMAGTACKSR 179  
QY 179 PLGSGDCPLELKNVLEDEGMDVDTGYGAMDEFTLDTCKEVPDLTQSCICKYPDYLOMS 238  
DB 180 APRPTDCEPLELKNVLEDEGMDVDTGYGAMDEFTLDTCKEVPDLTQSCICKYPDYLOMS 239  
QY 239 ADPYGDSMFCLREQLFARHFNRACTMGDTVPQSLYIKGT-GMARSPSCVYSPSPSGS 298  
DB 240 AEPIYGDSMFCLREQLFARHFNRACTMGDTVPQSLYIKGT-GMARSPSCVYSPSPSGS 299  
QY 300 MYSSDSOLFENKPYWLHKAQGHNNNGICWNLFTVVDTRSTNLTICASTQSPVPGQYDA 358  
DB 299 IVTSDSOLFENKPYWLHKAQGHNNNGICWNLFTVVDTRSTNLTICASTQSPVPGQYDA 358  
QY 359 TKFKQYSRHVEYDIQIFQLCTITLADVMSYIHSNNSSTLEDMNFGVPPPTSLVDY 418  
DB 359 TSFKQYSRHVEYDIQIFQLCTITLADVMSYIHSNNSSTLEDMNFGVPPPTSLVDY 418  
QY 419 YRFQSVATTCOKDAAPAKNDPYDKLFWNDLKEKESLDDOYPLGKFLVQAGLR 477  
DB 419 YRFQSVATTCOKDAAPAKNDPYDKLFWNDLKEKESLDDOYPLGKFLVQAGLR 478

QY 478 KPTIGPKRSAPSATTSSKPAKRRVR 504  
 DB 479 VSV--SRKRAAPSSTPTSSPATKRRKR 503

RESULT 14  
 ID V1\_HPV53 STANDARD: PRT; 499 AA.

DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Major capsid protein L1.

GN L1.  
 OS Human papillomavirus type 53.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OX NCBI\_TaxID=10619;

RP SEQUENCE FROM N.A.  
 RX MEDLINE=94265501; PubMed=8205838;  
 RA Delius H., Hofmann B.;  
 RT "Primer-directed sequencing of human papillomavirus types";  
 RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).

RP SEQUENCE OF 300-343 FROM N.A.  
 RX MEDLINE=92407963; PubMed=1326639;  
 RA Chan S.Y., Bernard H.O., Ong C.K., Chan S.P., Birgit H., Delius H.;  
 RT "Phylogenetic analysis of 48 papillomavirus types and 28 subtypes and  
 variants: a showcase for the molecular evolution of DNA viruses.";  
 RL J. Virol. 66:5714-5725(1992).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL; X74482; CAA52595.1; -  
 DR EMBL; M96298; AAA47037.1; -  
 DR PIR; S36531; S36531.  
 DR InterPro: IPR002210; PV\_capsid\_L1.  
 DR Pfam: PF00500; late\_protein\_L1; 1.  
 DR PRINTS: PR00865; HPVcapsid.L1.  
 DR ProDom: PD000544; PV\_capsid\_L1; 1.  
 DR Coat protein; late protein.  
 KW SEQUENCE 499 AA; 55722 MW; 5221961A3FDD5A66 CRC64;

Query Match 67.9%; Score 1855; DB 1; Length 499;  
 Best local Similarity 66.3%; Pred. No. 4.1e-153;  
 Matches 333; Conservative 78; Mismatches 87; Indels 4; Gaps 2;

QY 1 MALMRPSDNTVYLPPPSVARVNTDDYVTRTSIFYHAGSSRLTYVGNPFYRPVAGGAKNQ 60  
 DB 1 MAWMRPSDSKVTLLPPVSKVITTDAYVKTTFYHAGSSRLTYVGHPRYPYPSKSG--KA 58  
 QY 61 DIPKVSAYQYRVFVQLPDPNKFGLPDNSIYNPETQRLVWACAGVEIGRGPLGVLSGH 120  
 DB 59 DIPKVSAYQYRVFVQLPDPNKFGLPDNTIFNPDERLVWACVLEIGRGPLGVLSGH 118  
 QY 121 PFYNTKLDTESSHAATSNVSEVDVNDVYKOTQCLICGAPAIGEHNAKGTACKSRPL 180  
 DB 119 PFENRDLDTRESSSIADIDTAPDSRDVNPVDPKOTQCLICGAPAIGEHNTKGTACKSRPT 178  
 QY 181 SGGDCPPELTKNTVLEDDGMVDITGAMDFSTLQDTKCEVPLDIOQSTICKYDYLQMSAD 240  
 DB 179 TGDGCPPELTLNPSLJEDGDMVDITGALNFKALQESKSDVPLDIVQSTCKYDYLKMSAD 238  
 QY 241 PYGDSMFCLRREQLFARHFNRRAGTMDTVPOSILYIKGTGMRASPGSCVYSPSPSGSI 300  
 DB 239 AVGDGSMWFYLRREQLFTRHFNRRAGTMDTVPOSILYIKGTGMRASPGSPSYVATPSGSKI 298

QY 301 TSDSQLEPNRPYLHAKQGHNNIGICWHNQLFVTVVDTTSTNTLTICASTOSPPVGOYDARK 360  
 DB 299 TSEAQLFNKRPYWLQRAQGHNNIGICWHNQLFVTVVDTTSTNTLTICASTOS--MSTYNSHQ 356  
 QY 361 FKOYSRHVEYDLOFIFQCLITITLADWMSYHSHSNSSILIEOMNFVPPPTSLYDTR 420  
 DB 357 IKOYVRHAEVEYDLOFIFQCLITISLAEVAYLFTNMSTLLEDMNITLSPVATSLDCKR 416  
 QY 421 FVOSVAITCQKADAPAPENKDPYDKLFWMVNDLKEKESLDIDQYPLGRKFLYQGRKRP 480  
 DB 417 YVSAITCQKADAPAPENKDPYDKLFWMVNDLKEKESLDIDQYPLGRKFLYQGRKRP 476  
 QY 481 IGPKRSAPSATTSSKPAKRRVR 502  
 DB 477 VSSKRSASTTSTSPSSKRRK 498

RESULT 15  
 ID V1\_RHPV1 STANDARD: PRT; 501 AA.

DT 01-FEB-1991 (Rel. 19, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Major capsid protein L1.

GN L1.  
 OS Rhesus papillomavirus type 1 (Rhpv 1).  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OX NCBI\_TaxID=10570;

RP SEQUENCE FROM N.A.  
 RX MEDLINE=91135018; PubMed=1847267;  
 RA Ostrow R.S., Labresh K.V., Fares A.J.;  
 RT "Characterization of the complete Rhpv 1 genomic sequence and an  
 RL integration locus from a metastatic tumor";  
 RL Virology 181:424-429(1991).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL; M60184; AAA79318.1; -  
 DR PIR; H38503; P1WKL1.  
 DR InterPro: IPR002210; PV\_capsid\_L1.  
 DR Pfam: PF00500; late\_protein\_L1; 1.  
 DR PRINTS: PR00865; HPVcapsid.L1.  
 DR ProDom: PD000544; PV\_capsid\_L1; 1.  
 DR Coat protein; late protein.  
 KW SEQUENCE 501 AA; 55635 MW; 6E808DF130E440A CRC64;

Query Match 67.9%; Score 1855; DB 1; Length 501;  
 Best local Similarity 67.1%; Pred. No. 4.1e-153;  
 Matches 338; Conservative 66; Mismatches 96; Indels 4; Gaps 3;

QY 1 MALMRPSDNTVYLPPPSVARVNTDDYVTRTSIFYHAGSSRLTYVGNPFYRPVAGGAKNQ 60  
 DB 1 MAWMRPSDSKVTLLPPVSKVITTDAYVKTTFYHAGSSRLTYVGHPRYPYPSKSG--GNKV 59  
 QY 61 DIPKVSAYQYRVFVQLPDPNKFGLPDNSIYNPETQRLVWACAGVEIGRGPLGVLSGH 120  
 DB 60 SVPKVSGDQYRVFVQLPDPNKFGLPDNAYFDPNQLVWACVLEIGRGPLGVLSGH 119  
 QY 121 PFYNTKLDTESSHAATSNVSEVDVNDVYKOTQCLICGAPAIGEHNAKGTACKSRPL 180  
 DB 120 PLNTKLDTESSHAATSNVSEVDVNDVYKOTQCLICGAPAIGEHNAKGTACKSRPL 177  
 QY 181 SGGDCPPELTKNTVLEDDGMVDITGAMDFSTLQDTKCEVPLDIOQSTICKYDYLQMSAD 240

```

Db      : ||| ||| :||:|||||:|||||: || | :||:|||||:|||||:|:|
178 AAGDCPALIELVNSYIODGDWDTGYGAMDEFNALQANKSDVPIDICTSVCKYPDILKMASD 237
QY      PYGDSMFECLEHREQLFARHFWNRAGTMDTVPOSLYIKGTGMRASPGSCVYSPSPSGSIY 300
241 ||||:|||||:| |||:|||||:| ||||:| |||:| |||:| |||:|
Db      PYGDSLEFFYLKREQMFTYRHLFNRAGTMDSVDPDLTIKSGSNVKLASHVFTYFPGSGMV 297
QY      301 TSDSQLFNKPYYMLHKAQGHNGICWHNOLFVTVVDTTRSTNLITICASTQSPVPGQYDATK 360
Db      298 TSDAQLFNKPYWLQKAGHNGICWGNQVFLTVVDTTRSTNMTLCASTASTVTPYNNES 357
QY      361 FKQISRHVEEDLOFIFQLCTITITADVMSTIHSNNSILEDMMNFCVPPPTTSIVDTYR 420
Db      358 FKEXLRHVEEDLOFIFQLCKVTLNTEVMAYIHSMDASILEDMMNFGLOPPSGSLODTYR 417
QY      421 FVQSVAITCOKDAAPAEKNDPYDKLFKNVVDLKEKFSLDLDQYPLGRKFLYQAGLRRKPT 480
Db      418 FVTSNAITCOKPAPPEKEDEPLAKYTFWEVDLKEKFSADLDQFPLGRKFLYQAGRRAPT 477
QY      481 IGPRKRSAPSATYSSKPAKRVVR 504
Db      : ||:|||||:| |||:| |||:| |||:| |||:| |||:|
478 LRAPKRTA-SSTSSSPRKRRTK 500

```

Search completed: October 17, 2003, 10:59:44  
 Job time : 17.7128 secs



Qy	120	HPFYKKLDLDTSSHAATSNVSEEDVADNVSVKQYOLCLIGCAPAIGEMWAGTACKNSR	179
Db	147	HPFYKKLDLDTSSHAATSNVSEEDVADNVSVKQYOLCLIGCAPAIGEMWAGTACKNSR	206
Qy	180	LSQGDCEPLELKNITVLEDGDVNTGYGAMDFSTLQDTRKEVPLDIQSICKYPDYLOMSA	239
Db	207	LQPGDCPEPLELKNITIEDGDVNTGYGAMDFSTLQDTRKEVPLDIQSICKYPDYLOMSA	266
Qy	240	DPYGDSMFFCLRREOLFARHFWNRAIGMDIYPOSLYIKGT--GMRASPGSCYSSPSGS	297
Db	267	DPYGDSMFFCLRREOLFARHFWNRAIGMDIYPOSLYIKGT--GMRASPGSCYSSPSGS	326
Qy	298	SIVTSDSLQFNKPYWLHKRAGGNHNCICWNNOLFVYVYDTRSTNLTLCASTOSPVYQGDY	357
Db	327	SIVTSDSLQFNKPYWLHKRAGGNHNCICWNNOLFVYVYDTRSTNLTLCASTOSPVYQGDY	386
Qy	358	ATKFKOYSRHVEEYDLOEITFOLCTITLTPADVASYIHSNMSSILIEDNFCVPEPPTSLVD	417
Db	387	PTKFKFYSRHVEEYDLOEITFOLCTITLTPADVASYIHSNMSSILIEDNFCVPEPPTSLVD	446
Qy	418	TYRFVQSAVITQCKDAPAEKNQDPYDKLKFVNVNDLKEKTSLLDDQYPLGRKFLVQAGLR	477
Db	447	TYRFVQSAVITQCKDQPTPEKQDPYDKLKFVNVNDLKEKTSLLDDQYPLGRKFLVQAGLR	506
Qy	478	KPTIGRRKRSASATSSKPAKRVVYRRAK	507
Db	507	KPTIGRRKRSASATSSKPAKRVVYRRAK	536

	RESULT 2			
ID	Q9WHG7	PRELIMINARY;	PRF:	504 AA.
AC	Q9WHG7/;			
DT	01-NOV-1999 (TREMBLrel. 12, Created)			
DT	01-NOV-1999 (TREMBLrel. 12, Last sequence update)			
DE	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	Putative major capsid protein l1.			
CN	l1.			
OS	Human papillomavirus candidHPv85.			
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;			
OC	Papillomavirus.			
OX	NCBITaxID=151757;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20047972; PubMed=10580054;			
RA	Chow V.T.R., Leong P.W.F.;			
RT	*Complete nucleotide sequence, genomic organization and phylogenetic			
RT	analysis of a novel genital human papillomavirus type, HLT7474-S.";			
RL	J. Gen. Virol. 80:2923-2929(1999).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Chow V.T.R., Leong W.F.;			
RL	Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.			
DR	EMBL; AF131950; AAD24188.1; -			
DR	InterPro; IPR002210; PV_capsid.f1.			
DR	Pfam; PF00500; late_protein_l1.1			
DR	PRINTS; PR00865; HPVcapsIDL1.			
DR	PRODOM; PD000544; PV_capsid.L1.			
SQ	SEQUENCE 504 AA; 56435 MW; 3852CE9203E3D839 CRC64;			
Query Match				
Best Local Similarity 81.4%; Score 2223.5; DB 12; Length 504;				
Matches 400; Conservative 47; Mismatches 54; Indels 1; Gaps 1.				
QY	1 MALMRPSDNRYLLPPPSVARYVNNTDDYTRTSIFYHAGSSRLLTGNPFYRVPAAGGNKQ	60		
Dd	1 MALMRRSSDSKYLLPSPSAKAVNNDDVTTRTSTFYHAASSRLLYGHNYKYKTSNGGRKQ	60		
QY	61 DIPRYVASQYRVEFRVQLDPDNKKFGSLPDNSIYNPETQRLLVMACAGYEIGRGOPGLGYLSGH	120		
Dd	61 DIPKYSAYQYRVEFRVTLDPDNKKFGSLPDNTINYPETQRLLVMACVGMEVRGQPLGYLSGH	120		
QY	121 PFYKKLDFTESHAAITSNVSDVDVNDVSKQYDQLCLGCAPALGEHWMAQTACKSRPL	180		

Db	121	PFYFNKLDIDPENSIVASVYTHDTRDQVSVDYQIQTOLCIGCPALGEMMAGTACKGAV	180
QY	181	SGQDCCPPELEKNTVLEDGMDVTGCGAMDEFTLQDTKCEVPDLTQISICTKPYDLYQMSAD	240
Db	181	QTDCCPPELELVNYPIDGDMIDTGYAMDFSTLQONKKESEVPDLTQISICTKPYDLYQMSAD	240
QY	241	PYGDSMPEFCRLRRQLEFARHFNMWNRAGMGTVPQSLYXIGTGGRASPSCVYSPSPSGSIV	300
Db	241	AYGDSMPEFCRLRRQLEFARHFNMWNRAGTIGDAVFETIYXIGTDNRATPSCSISPSPSGSMV	300
QY	301	TSDSQLEFNKPYWLHKAQAGNNGICWHNQLFVTVDTRSTNLTJCASTQSPVQOYDATK	360
Db	301	SSDAQOFNKPYWLHKAQAGNNGICWHNQLFTVVDTRSTNLTJCASTQSPVQOYDATK	360
QY	361	FKQYSRHHVEYDQLEFQCLCTITLADVSYSHSNSSLIIDMNGVGPPTSLVNTYR	420
Db	361	FKETTRHVEYDQLEFQCLCTITLTDVSYSHNNDPTLILSMNGVSPPSASLVDTYR	420
QY	421	FVQSVAILTCQDAAPAEKNDPYDKLEFNVNVDLEKFFSLDDQYPLGRKFLVQAGILRRKPT	480
Db	421	FLQSSAITCQDDVYVPOKQDPEYKLEFNVNVDLEKFFSSDDQYPLGRKFLVQAGILRRKPT	480
QY	481	IGPRKRSABPATSSSKPAKRVR	502
Db	481	IGPRKRVRA-STSTATPRSKRR	501

```

RESULT 3
ID      081971          PRELIMINARY;          PRT;       508 AA.
AC      081971;
DC      01-NOV-1996 (TREMBLrel. 01, Created)
DT      01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DR      01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE      ORF putative I1 protein.
OS      Human papillomavirus type 59..
OC      Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OX      NCBI_TaxID=37115;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94303229; PubMed=8030272;
RA      Rho J., Roy-Burman A., Kim H., de Villiers E.M., Matsukura T.,
RA      Choe J.;
RT      "Nucleotide sequence and phylogenetic classification of human
RL      papillomavirus type 59." ;
RN      Virology 203:158-161(1994).
RP      [2]
RN      SEQUENCE FROM N.A.
RA      Choe J.;
RL      Submitted (FEB-1994) to the EMBL/GenBank/DDBJ databases.
DR      EMBL; X77858; CA54856.1; -.
DR      InterPro; IPR002210; PV_capsid_I1.
DR      Pfam; PF00500; late_protein_I1.
DR      PRINTS; PR00865; HPVcapsI1.
DR      PRODOM; PD000544; PV_capsid_I1; 1.
SQ      Sequence   508 AA; 56841 MW; F89A2D8C009EA5BA CRC64;

Query Match           81.4%; Score 2223.5; DB 12; Length 508;
Best Local Similarity 79.4%; Pred. No. 2e-191;
Matches 400; Conservative 49; Mismatches 54; Indels 1; Gaps 14.

Oy      1 MALWRPSNTYLLPPSPVARVVNDYDTRTSIFPHAGSSRLTLVGNFYRVPAGGKQ 60
Db      1 MALMRSSDNKKYLDPSPSAKYVSIDEYTKTISITFHAGSSRLTLVGHFYRVPKGNGRQ 60
Oy      61 DIPRYSAQYRVFRVOLEDPDKFKGLPDNSIYNPFTQRLVMWACAGVEIGRGOPLAGYLSGH 120
Db      61 DVPRYSAYQYRVERFKLLDDPKFKGLPDVTYVDPNSGRLVMWACGVGEIRGQPLGYLSGH 120
Oy      121 PFYKKLDPTESHAAITSNVSDYVDYDNNVSVDYRKQQLCLGCAPALGEHMAAGTGACKSRPL 180
Db      121 PFYKKLDPTESHAAITSNVSDYVDYDNNVSVDYRKQQLCLGCAPALGEHMAAGTGACKSRPL 180

```





QY	298	SIYVSDSOLFKKPWLKKAQGHNNIGCMHNOLEFVYVDTSTNTLTCASQSPVPGQYD	357
Db	300	SMATSEQOLNKKPWLKRAQGHNNIGCMANOLEFVYVDTSTNTMTLCVSTET--SAYTD	357
QY	358	ATKFKQYSRHYEEDLOFIFQOLCTITLTADVMYSIHSNMSILEDNMNFGVPPPTTSYLD	417
Db	358	ATKFKETLRHGEEDLOFIFQOLCTVTLPELMAYLHMNTSLLEDNMNGLLPPSTSLD	417
QY	418	TYRVSQVAITCCKDAAPAEKKDPYDKLKEFMNVDLKEKFSJDLDDYPLGKRFVQAGLR	477
Db	418	TYRFLTSSAITCQADAPPECKODEPYAKLNFWDVLKORFSDLSQFLGRKFLMAGVGT	477
QY	478	KPTIGPKRRSAPASATSSKPAKRVYRKR	507
Db	478	RSSISVRRK--SATTTSRTAAAKRKRTK	504
RESULT 6			
	081958	PRELIMINARY:	PR1: 503 AA.
AC	Q81958		
DT	01-NOV-1996	(TREMBLrel. 01, Created)	
DT	01-NOV-1996	(TREMBLrel. 01, Last sequence update)	
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)	
DE		late protein (Fragment).	
GN	L1.		
OS	Human papillomavirus type 10.		
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;		
OC	Papillomavirus.		
OX	NCBI_Taxid=10603;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=9426501; PubMed=8205838;		
RA	DeJius H., Hofmann B.;		
RT	"Primer-directed sequencing of human papillomavirus types.";		
RL	Cur. Top. Microbiol. Immunol. 186:13-31(1994).		
DR	EMBL: X74465; CAAS2495.1; -;		
DR	InterPro: IPR002210; PY_capsid_L1.		
DR	Pfam: Pf00500; late_protein_L1; 1.		
DR	PRINTS: PR00865; HPVcapsIDL1.		
DR	ProDom: PD00544; PY_capsid_L1; 1.		
FT	NON TER		
SQ	SEQUENCE 503 AA: 56041 MW: 339739258A79E2A CRC64:		
Query Match			
	Best Local Similarity	69.8%; Score 1909; DB 12; Length 503;	
	Matches 342; Conservative 70; Mismatches 90; Indels 2; Gaps 2;		
QY	1	MALMRSPDNNYYLPPPSVARYNNDDVYTRISIFTHAGSSRLTYGNRYFRVAPAGGKKQ	60
Db	1	MALMRSSDNNLYLPPPTPSKVLSTDYDVTIRNITYAGTSRLTYGVHYPFPIRNSNNKV	60
QY	61	DIPKVSAYQYKVFQVLPDPNKFGLPDSIYNPETORLWMAAGVEIGROPGLGSLGH	120
Db	61	DVPRVSAFYQRYEVRRLPDPNKFGLPDAIRIYNPAERLYMACTGEVARGGPLGVLSGH	120
QY	121	PYINKLDTFSSHAATSNVSEVDVBNVSVYDKQFQLCGLCAPAIGEHMACTACKSRPL	180
Db	121	PLYNKLDTFENSINAHGPICGDSHDNISVDNKKQFQLCTIGCTPPMGHEMGKTPCRNPP	179
QY	181	SOGGCPPELLEKNYLYLEDGGDMWDTCYGAMDEFTDITCKEYVLDITQSTICKTPDIQMSAD	240
Db	180	AQGGCPPELLELTSYLDGGDMWDTYGADEFDAIDLMSNDVIDIDICQSTCKPDIYLGMAE	239
QY	241	PYGDSMPECLREDFARHFENNRAGTGMGVPOSLYIKGTGMRASPGSCYSPSPSGSIV	300
Db	240	PYGDSMPEFYLRREDFARHFENNRASAVGDAIPDPFIKSNRGAGDVGSAYISPPRSGSMV	299
QY	301	TSDSQLEFKPEYWLKKAQGHNNIGCMHNOLEFVYVDTSTNTLTCASQSPVPGQYDATK	360
Db	300	TSEQLFENKPEYWLKRAQGHNNIGCMANOLEFVYVDTSTNTMCLCVPEASPAITYDATK	359
QY	361	FKQYSRHYEEDLOFIFQOLCTITLTADVMYSIHSNMSILEDNMNFGVPPPTTSYLDYTR	420

[illegible]

```

Db      372 YSSTFEKELNTHETEDLOFIFOLCKITHTPEIMSTLHNMNPTLIDENMFVIPPSTSL 431
Qy      416 VDTYFVQSVATTCOK--DAAPAEKNDPYDKLFKMNVDLEKEFSLDDQYPLGRKFLVQAG 474
Db      432 DDTYRFLNRSALTCKCKGTAPAEKPKDPYKLSFMDYDLKERLSTLDQPLGRKFLQAG 491
Qy      475 LRRKPTIGPKRSAPSATSS--KPAKR 500
Db      492 GR--PSSVPRKRAAPVSTSKAPRKRAK 518

```

## RESULT 8

```

Q8B5X1  PRELIMINARY; PRT; 525 AA.
AC      Q8B5X1;
DT      01-MAR-2003 (TREMBLrel. 23, Created)
DT      01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DE      01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE      Putative major capsid protein L1.
GN      L1.
OS      Human papillomavirus type 54.
OC      Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC      Papillomavirus.
OX      NCBI_Taxid=37113;
RN      (1)
RP      SEQUENCE FROM N.A.
RA      Terai M., Burk R.D.;
RT      "Cervical HPV in Evolution: Genomic sequence of AE9, a subtype of
RT      HPV24."
RL      Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF436129; AA015434.1; -.
SQ      SEQUENCE 525 AA; 58773 MW; 2BBE2F8A3DEC869B CRC64;

```

Query Match 68.1%; Score 1862.5; DB 12; Length 525;  
 Best local Similarity 69.0%; Pred. No. 6.7e-159;  
 Matches 351; Conservative 57; Mismatches 84; Indels 17; Gaps 7;

```

Qy      1 MALMRPSDNTVYLLPPPSVAVVNTDDYVTRTSIFYHAGSSRLTYGNPFYRPVAGGKMQ 60
Db      27 MAMMRPSERKXYLLPPTPVSKVSTDEYVTRTSIYHASSRLLAVGHPYFKYOK--STNKO 85
Qy      61 DIPKSAVQYRVFVQVLPDPNFKGLPDNSIYNPETQRLVWACAGVEIGRGQPLGVLSGH 120
Db      86 SLPKSAVQYRVFVQVLPDPNFKGLPDNSIYNPETQRLVWACAGVEIGRGQPLGVLSGH 145
Qy      121 PFYKRLDPTSS-----HAATSNVSEVDYRDNVSVDKOTQCLTIGCAPAIGEMHAKGTACK 176
Db      146 PLANKLIDTENAKPKVVGAGTDN-----KENVSMQKQVQLCLIGCTPPIGEMHAKGNLCT 200
Qy      177 SRPLSQGDCPPLLEKNTVLEDGMDVDTGYGAMDFSTLDPTKCEVPLDIOQSICTYDYLQ 236
Db      201 PNTLAGDCPPLLEKNTVLEDGMDVDTGYGAMDFSTLDPTKCEVPLDIOQSICTYDYLQ 260
Qy      237 MSADYDGSMPFCLAREQLFAHFHNNRAGTMDGTVPQSLYIKGTGMRASPGCYVSPSPS 296
Db      261 MAAEYIGSILFYLRRQEFVNHMLNRAGTMEPEVPNDLYIKSS--GAPDSIYAAPPS 318
Qy      297 GSIVTSDQLFNKPYMLHKAQGHNGICMHNOLFVTVVDTTRSTMLTICASQSPVPGQY 356
Db      319 GSMVTSYQIFNKPWMLRAQGGNNGICMGNOLFVTVVDTTRSTMLTICATATS--QDTF 376
Qy      357 DATKQYSRHVEEYDLQIFOLCTITTLADVMSYIHSNNSILEDWNEGVPPPTSLV 416
Db      377 NNANKEFYIRHVEEYDLQIFOLCTITTLADVMSYIHSNNSILEDWNEGVPPPTSLV 436
Qy      417 DTYRFGVQSVATTCOKDAAP--ENKDPYDKLFKMNVDLEKEFSLDDQYPLGRKFLVQAGL 475
Db      437 DTYRFGVQSVATTCOKDAAP--ENKDPYDKLFKMNVDLEKEFSLDDQYPLGRKFLVQAGL 496
Qy      476 RRRKPTIGPKRSAPSAT--TSSKPAKRYR 502
Db      497 RRRPLRLRPVKRAAPSSSGSTSKRAKTRK 525

```

## RESULT 9

```

Q993Z3  PRELIMINARY; PRT; 503 AA.
AC      Q993Z3;
DT      01-JUN-2001 (TREMBLrel. 17, Created)
DT      01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE      01-OCT-2001 (TREMBLrel. 18, Last annotation update)
DE      Putative major capsid protein L1.
GN      L1.
OS      Human papillomavirus type 82.
OC      Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC      Papillomavirus.
OX      NCBI_Taxid=129724;
RN      (1)
RP      SEQUENCE FROM N.A.
RA      Terai M., Burk R.D.;
RT      "Cervical HPV in Evolution: Genomic Sequence of IS39/AE2, a Subtype
RT      of Oncogenic HPV 82 (W13B).";
RL      Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF293961; AKK28456.1; -.
DR      InterPro: IPR002210; PV_capsid_L1.
DR      Pfam: PF00500; Late_protein_L1;
DR      PRINTS: PR00865; HPVcapsidL1.
DR      ProDom: PD000544; PV_capsid_L1; 1.
SQ      SEQUENCE 503 AA; 56121 MW; F6C4F5D72BFCE93 CRC64;

```

Query Match 68.1%; Score 1861; DB 12; Length 503;  
 Best local Similarity 66.7%; Pred. No. 8.6e-159;  
 Matches 337; Conservative 74; Mismatches 90; Indels 4; Gaps 4;

```

Qy      1 MALMRPSDNTVYLLPPPSVAVVNTDDYVTRTSIFYHAGSSRLTYGNPFYRPVAGGKMQ 60
Db      1 MALMRPSDNTVYLLPPPSVAVVNTDDYVTRTSIFYHAGSSRLTYGNPFYRPVAGGKMQ 59
Qy      61 DIPKSAVQYRVFVQVLPDPNFKGLPDNSIYNPETQRLVWACAGVEIGRGQPLGVLSGH 120
Db      60 EIPKSAVQYRVFVQVLPDPNFKGLPDNSIYNPETQRLVWACAGVEIGRGQPLGVLSGH 119
Qy      121 PFYKRLDPTSS-----HAATSNVSEVDYRDNVSVDKOTQCLTIGCAPAIGEMHAKGTACKSRL 180
Db      120 PLNKTYDDTENSRRVANGVQDVNDISVDNKOTQCLTIGCAPAIGEMHAKGTACKSNPV 179
Qy      181 SQGDCPPLLEKNTVLEDGMDVDTGYGAMDFSTLDPTKCEVPLDIOQSICTYDYLOMSAD 240
Db      180 PGDCPPLLEKNTVLEDGMDVDTGYGAMDFSTLDPTKCEVPLDIOQSICTYDYLOMSAD 239
Qy      241 PYGDSMFCLAREQLFAHFHNNRAGTMDGTVPQSLYIKGTGMRASP--GSCVYSPSPSGSI 299
Db      240 TYGNSMFFELRREQIFARHYNRAGVGDALPDKAYIKGTGTRDPISGYSATPPSGSM 299
Qy      300 VTSDSQLEFNKPYMLHKAQGHNGICMHNOLFVTVVDTTRSTMLTICASQSPVPGQYDAT 359
Db      300 VTSDSQLEFNKPYMLHKAQGHNGICMHNOLFVTVVDTTRSTMLTICATPSVAQTPTPT 359
Qy      360 KEKQYSRHVEEYDLQIFOLCTITTLADVMSYIHSNNSILEDWNEGVPPPTSLVDTY 419
Db      360 NKQYIRHVEEYDLQIFOLCTITTLTEVMALHNDSTILEQWNGFLTPPSALEDAY 419
Qy      420 RYVQSVATTCOKDAAPAEKNDPYDKLFKMNVDLEKEFSLDDQYPLGRKFLVQAGLRRKP 479
Db      420 RYVQSVATTCOKDAAPAEKNDPYDKLFKMNVDLEKEFSLDDQYPLGRKFLVQAGLRRKP 479
Qy      480 TIGPKRSAPSATSSKPAKRYR 504
Db      480 RRG--LRRPAPSSSSSS--AKRRKVK 502

```

## RESULT 10

```

Q91RS2  PRELIMINARY; PRT; 503 AA.
AC      Q91RS2;

```

DF 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)  
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)  
DE DNA, complete genome.  
GN L1.  
OS Human papillomavirus type 82.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=129724;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Sata T., Matsukura T.;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MEDLINE-20087389; PubMed-10618284;  
RX Kuno N., Sata T., Sato Y., Sugase M., Matsukura T.;  
RT "Molecular cloning and nucleotide sequence analysis of a novel human  
papillomavirus (type 82) associated with vaginal intraepithelial  
neoplasia";  
RT Clin. Diagn. Immunol. 7:91-95(2000).  
RL EMBL; AB027021; BAA90742.1; -  
DR InterPro; IPR002210; PV\_capsid.L1.  
DR Pfam; PF00500; late\_protein.L1; 1.  
DR PRINTS; PR00865; HPVcapsid.L1.  
DR PRODOM; PD000544; PV\_capsid.L1; 1.  
SQ SEQUENCE 503 AA; 56209 MW; 526E427B3A2C6CC3 CRC64;

Query Match 67.8%; Score 1852; DB 12; Length 503;  
Best Local Similarity 66.1%; Pred. No. 5.5e-158;  
Matches 334; Conservative 75; Mismatches 92; Indels 4; Gaps 4;

QY 1 MALMRPESDNTVYLPPEPVANVTDDYVTRTSTFYHAGSSRLTLVGNPFRVPAAGNKQ 60  
DB 1 MALMRKTDSDGVYLPPEPVANVTDDYVTRTSTFYHAGSSRLTLVGNPFRVPAAGNKQ 59  
QY 61 DIPKVSAYGYRVFVQLPDNPKFGLPDNSIYNPETORLVMACAGVEIGRQPLGVSGH 120  
DB 60 EIRPVSAFOYRVFVQLPDNPKFGLPDNPKFGLPDNPKFGLPDNPKFGLPDNPKFGL 119  
QY 121 PFYKLDLDTSSHAATSNVSEVDVNDVYKOTQCLICAPALIGEMAKGTACKSRPL 180  
DB 120 PLFKKYDTEHSRFRANQDQVDVNDVYKOTQCLICAPALIGEMAKGTACKSRPL 179  
QY 181 SQGGCPPELEKNTVLEGGDWDVGYGAMDFSTLDTCEVPLDLCOSICTKPYDLOASD 240  
DB 180 PQGGCPPELEKNTVLEGGDWDVGYGAMDFSTLDTCEVPLDLCOSICTKPYDLOASD 239  
QY 241 PYGDSNFFCLREQLFARHFWNRAGTMDVPOSLYIKGTGMRASPGSCVSPSPSGSI 299  
DB 240 TYGSMFPLRREIFARHFWNRAGTMDVPOSLYIKGTGMRASPGSCVSPSPSGSI 299  
QY 300 VTSDSQLFNKPYWLHKAQGNNGICWNNQLEFVTVVDTTRSTNLTICASTOSPVPQOYDAT 359  
DB 300 ITSOSQFNKPYWLHKAQGNNGICWNNQLEFVTVVDTTRSTNLTICASTOSPVPQOYDAT 359  
QY 360 KFKQYSHRVEYDLOFIFOLCTITLADVMSYIHSNMSSTLEDMNFGVPPPTSLYDTY 419  
DB 360 NFQOYIRHGEYDLOFIFOLCTITLADVMSYIHSNMSSTLEDMNFGVPPPTSLYDTY 419  
QY 420 RFVQSVATCOKDAAPAEKNKDPYDKLFKFWNVDLKEKPSLDLDQYPLGRKFLYQAGLRKP 479  
DB 420 RFVQSVATCOKDAAPAEKNKDPYDKLFKFWNVDLKEKPSLDLDQYPLGRKFLYQAGLRKP 479  
QY 480 TIGPRKRSAPASATTSKPAKRVRYR 504  
DB 480 TIGPRKRSAPASATTSKPAKRVRYR 504  
QY 480 RPKRPAPSSSSASSS-AKRRKRYK 502  
DB 480 RPKRPAPSSSSASSS-AKRRKRYK 502

RESULT 11  
Q99FW7 PRELIMINARY; PRT; 503 AA.  
ID Q99FW7  
AC Q99FW7;

DF 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-OCT-2001 (Tremblrel. 18, Last annotation update)  
DE Putative major capsid protein L1.  
GN L1.  
OS Human papillomavirus type 84.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=150546;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE-21066735; PubMed-11145894;  
RX Teraï M., Burk R.D.;  
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Burk R.D., Teraï M.;  
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.  
RX EMBL; AF293960; AAK09277.1; -  
DR InterPro; IPR002210; PV\_capsid.L1.  
DR Pfam; PF00500; late\_protein.L1; 1.  
DR PRINTS; PR00865; HPVcapsid.L1.  
DR PRODOM; PD000544; PV\_capsid.L1; 1.  
SQ SEQUENCE 503 AA; 56438 MW; 1AD07CE3C1BFC0E4 CRC64;

Query Match 67.2%; Score 1836; DB 12; Length 503;  
Best Local Similarity 66.4%; Pred. No. 1.5e-156;  
Matches 340; Conservative 69; Mismatches 89; Indels 14; Gaps 7;

QY 1 MALMRPESDNTVYLPPEPVANVTDDYVTRTSTFYHAGSSRLTLVGNPFRVPA--GGG 57  
DB 1 MALMRPESDNTVYLPPEPVANVTDDYVTRTSTFYHAGSSRLTLVGNPFRVPA--GGG 56  
QY 58 NKO-DIPKVSAYGYRVFVQLPDNPKFGLPDNSIYNPETORLVMACAGVEIGRQPLGVSGH 116  
DB 61 NKRAITPKVSGYQYRVFVQLPDNPKFGLPDNPKFGLPDNPKFGLPDNPKFGLPDNPKFGL 120  
QY 117 LSGHFFYKLDLDTSSHAATSNVSEVDVNDVYKOTQCLICAPALIGEMAKGTACKSRPL 176  
DB 121 TSQGPLYKLDLDTSSHAATSNVSEVDVNDVYKOTQCLICAPALIGEMAKGTACKSRPL 179  
QY 177 SRPLSDGCPPELEKNTVLEGGDWDVGYGAMDFSTLDTCEVPLDLCOSICTKPYDLOASD 236  
DB 180 NVQYRAGDCPPLEKNTVLEGGDWDVGYGAMDFSTLDTCEVPLDLCOSICTKPYDLOASD 239  
QY 237 MSADPYGDSNFFCLREQLFARHFWNRAGTMDVPOSLYIKGTGMRASPGSCVSPSPSGSI 296  
DB 240 MAEPYDGCNFFCLREQLFARHFWNRAGTMDVPOSLYIKGTGMRASPGSCVSPSPSGSI 299  
QY 297 GSIYVTSDSQLFNKPYWLHKAQGNNGICWNNQLEFVTVVDTTRSTNLTICASTOSPVPQOYDAT 356  
DB 300 GSIYVTSDSQLFNKPYWLHKAQGNNGICWNNQLEFVTVVDTTRSTNLTICASTOSPVPQOYDAT 357  
QY 357 DATKFOYSHRVEYDLOFIFOLCTITLADVMSYIHSNMSSTLEDMNFGVPPPTSLYDTY 416  
DB 358 KPTNEKKEIRHGEYDLOFIFOLCTITLADVMSYIHSNMSSTLEDMNFGVPPPTSLYDTY 417  
QY 417 DTYRFVQSVATCOKDAAPAEKNKDPYDKLFKFWNVDLKEKPSLDLDQYPLGRKFLYQAGLRKP 475  
DB 418 DTYRFVQSVATCOKDAAPAEKNKDPYDKLFKFWNVDLKEKPSLDLDQYPLGRKFLYQAGLRKP 477  
QY 476 RRRPTIGPRKRSAPASATTSKPAKRVRYR 507  
DB 478 R--STLVSRKRRTASAGT---PPASRRKRAKK 503

RESULT 12  
Q81017 PRELIMINARY; PRT; 504 AA.  
ID Q81017  
AC Q81017;  
DT 01-NOV-1996 (Tremblrel. 01, Created)



RA Icenogle J.P., Clancy K.A., Lin S.Y.;  
 RT "Sequence variation in the capsid protein genes of human  
 RT Papillomavirus type 16 and type 31.";  
 RL Virology 214:664-669(1995).  
 DR EMBL: U37217; AAA92892.1; -;  
 DR InterPro: IPR002210; PV\_capsid\_L1.  
 DR Pfam: PF00500; Late\_protein\_L1;  
 DR PRINTS: PR00865; HPVcapsidL1.  
 DR PRODOM: PD000544; PV\_capsid\_L1; 1.  
 DR SEQUENCE 505 AA; 56344 MW; 52D2D06922836F68 CRC64;

Query Match 66.9%; Score 1828; DB 12; Length 505;  
 Best Local Similarity 66.1%; Pred. No. 8e-156;  
 Matches 332; Conservative 69; Mismatches 97; Indels 4; Gaps 3;

```

QY 1 MALRRPSDNTVYLLPPPSVARYVNTDDYVTRTSIFYHAGSSRLTVGNPFYFVPAAGGNKQ 60
    1 MSLWLPSEATVYLLPPVPVSKVSTDEVYARNTIYHAGTSRLVAVGHYFFIKKPNNNKI 60
DB 1 MSLWLPSEATVYLLPPVPVSKVSTDEVYARNTIYHAGTSRLVAVGHYFFIKKPNNNKI 60
QY 61 DIPKVSAYQYRVFVQVLPDPNKKFGLPDSNIYNPEFQRLVMACAGVEIGRQPLGVGSGH 120
    61 LVPKVSGLQYRVFRIYLPDPNKKFGLPDSNIYNPEFQRLVMACAGVEIGRQPLGVGSGH 120
DB 61 LVPKVSGLQYRVFRIYLPDPNKKFGLPDSNIYNPEFQRLVMACAGVEIGRQPLGVGSGH 120
QY 121 PFYKKLDTESSHAATSVSEVDVNDVSDYKQVQLCLGCAPAIGEHMAGTACKSRPL 180
    121 PLNKKLDTEENASAYANAGVDNRECIISMDYKQVQLCLGCKPPIGEHMGSGSCNNVAV 180
DB 121 PLNKKLDTEENASAYANAGVDNRECIISMDYKQVQLCLGCKPPIGEHMGSGSCNNVAV 180
QY 181 SGGDCPPLLEKNTVLEDDVDTGYGAMDFSTLDTRCEVPLDIOQSIQCKYPDYLOMSAD 240
    181 TPGDCPPLLEINTVYIOGDMVDTFGAMDFSTLDTRCEVPLDIOQSIQCKYPDYLOMSAD 240
DB 181 TPGDCPPLLEINTVYIOGDMVDTFGAMDFSTLDTRCEVPLDIOQSIQCKYPDYLOMSAD 240
QY 241 PYGSMFFCLAREQLFARHFNWNRAGTMDYVPOSLYIKGTGMRASPGSCYVSPSGSIV 300
    241 PYGSLFFYLLRREDMFVHLEFNRAAGVENPDDLXIKSGPTANLASSNFFPPSGSMV 300
DB 241 PYGSLFFYLLRREDMFVHLEFNRAAGVENPDDLXIKSGPTANLASSNFFPPSGSMV 300
QY 301 TSDSOLFNRKPYWLKRAOGHNNGICMHNOLFVTVVDTTRSTNLITICASTQSPVPGQYDATK 360
    301 TSDAQIFNRKPYWLQRAOGHNNGICMGNOLFVTVVDTTRSTNLSICAAISTSEP-TYKNTN 359
DB 301 TSDAQIFNRKPYWLQRAOGHNNGICMGNOLFVTVVDTTRSTNLSICAAISTSEP-TYKNTN 359
QY 361 FKQYSRHVEEYDLOFIFQLCTITLADVMYSIHMSNSILEDWMFVGPPTTSVLDTYR 420
    360 FKEYLRHGEEDLOFIFQLCTITLADVMYIHSNSTILEDWMNGLOPPGGTLEDITYR 419
DB 360 FKEYLRHGEEDLOFIFQLCTITLADVMYIHSNSTILEDWMNGLOPPGGTLEDITYR 419
QY 421 FVQSVATTCCKDAAPAEKMDPYDKLKFVNVDLKEKFSLDLQYPLGRKFLVQAGLRKRP- 479
    420 FVTSQALACQKHTPPAREDEPLKRYTFWEVNLKEKFSADLDQFLGRKFLVQAGLRKRP 479
DB 420 FVTSQALACQKHTPPAREDEPLKRYTFWEVNLKEKFSADLDQFLGRKFLVQAGLRKRP 479
QY 480 -TIGPRKRSAPSATTSSKPAKR 500
    480 FTLGKRRK-ATPTTSTSTTAKR 500
DB 480 FTLGKRRK-ATPTTSTSTTAKR 500

```

## RESULT 15

08B5N5 PRELIMINARY; PRT: 531 AA.

AC 08B5N5.  
 DT 01-MAR-2003 (TREMBLrel. 23. Created)  
 DT 01-MAR-2003 (TREMBLrel. 23. Last sequence update)  
 DT 01-MAR-2003 (TREMBLrel. 23. Last annotation update)  
 DE Putative major capsid protein L1.  
 GN L1.  
 OS Human papillomavirus type 16.  
 OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
 OC Papillomavirus.  
 OC NCBI\_TaxID=10581;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Terai M., Ma Z., Burk R.D.;  
 RT "Cloning and Sequencing of Non-European Human Papillomavirus (HPV)  
 RT Variants Complete Genomes from Patients with Cervical Cancer by an  
 RT Overlapping PCR Method.";  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF472509; AAO15712.1; -.

SQ SEQUENCE 531 AA; 59541 MW; F1C433C8ADA313AA CRC64;  
 Query Match 66.9%; Score 1828; DB 12; Length 531;  
 Best Local Similarity 66.1%; Pred. No. 8.6e-156;  
 Matches 332; Conservative 69; Mismatches 97; Indels 4; Gaps 3;

```

QY 1 MALRRPSDNTVYLLPPPSVARYVNTDDYVTRTSIFYHAGSSRLTVGNPFYFVPAAGGNKQ 60
    27 MSLWLPSEATVYLLPPVPVSKVSTDEVYARNTIYHAGTSRLVAVGHYFFIKKPNNNKI 86
DB 1 MSLWLPSEATVYLLPPVPVSKVSTDEVYARNTIYHAGTSRLVAVGHYFFIKKPNNNKI 86
QY 61 DIPKVSAYQYRVFVQVLPDPNKKFGLPDSNIYNPEFQRLVMACAGVEIGRQPLGVGSGH 120
    87 LVPKVSGLQYRVFRIYLPDPNKKFGLPDSNIYNPEFQRLVMACAGVEIGRQPLGVGSGH 146
DB 87 LVPKVSGLQYRVFRIYLPDPNKKFGLPDSNIYNPEFQRLVMACAGVEIGRQPLGVGSGH 146
QY 121 PFYKKLDTESSHAATSVSEVDVNDVSDYKQVQLCLGCAPAIGEHMAGTACKSRPL 180
    147 PLNKKLDTEENASAYANAGVDNRECIISMDYKQVQLCLGCKPPIGEHMGSGSCNNVAV 206
DB 147 PLNKKLDTEENASAYANAGVDNRECIISMDYKQVQLCLGCKPPIGEHMGSGSCNNVAV 206
QY 181 SGGDCPPLLEKNTVLEDDVDTGYGAMDFSTLDTRCEVPLDIOQSIQCKYPDYLOMSAD 240
    207 TPGDCPPLLEINTVYIOGDMVDTFGAMDFSTLDTRCEVPLDIOQSIQCKYPDYLOMSAD 266
DB 207 TPGDCPPLLEINTVYIOGDMVDTFGAMDFSTLDTRCEVPLDIOQSIQCKYPDYLOMSAD 266
QY 241 PYGSMFFCLAREQLFARHFNWNRAGTMDYVPOSLYIKGTGMRASPGSCYVSPSGSIV 300
    267 PYGSLFFYLLRREDMFVHLEFNRAAGVENPDDLXIKSGPTANLASSNFFPPSGSMV 326
DB 267 PYGSLFFYLLRREDMFVHLEFNRAAGVENPDDLXIKSGPTANLASSNFFPPSGSMV 326
QY 301 TSDSOLFNRKPYWLKRAOGHNNGICMHNOLFVTVVDTTRSTNLITICASTQSPVPGQYDATK 360
    327 TSDAQIFNRKPYWLQRAOGHNNGICMGNOLFVTVVDTTRSTNLSICAAISTSEP-TYKNTN 385
DB 327 TSDAQIFNRKPYWLQRAOGHNNGICMGNOLFVTVVDTTRSTNLSICAAISTSEP-TYKNTN 385
QY 361 FKQYSRHVEEYDLOFIFQLCTITLADVMYSIHMSNSILEDWMFVGPPTTSVLDTYR 420
    386 FKEYLRHGEEDLOFIFQLCTITLADVMYIHSNSTILEDWMNGLOPPGGTLEDITYR 445
DB 386 FKEYLRHGEEDLOFIFQLCTITLADVMYIHSNSTILEDWMNGLOPPGGTLEDITYR 445
QY 421 FVQSVATTCCKDAAPAEKMDPYDKLKFVNVDLKEKFSLDLQYPLGRKFLVQAGLRKRP- 479
    446 FVTSQALACQKHTPPAREDEPLKRYTFWEVNLKEKFSADLDQFLGRKFLVQAGLRKRP 505
DB 446 FVTSQALACQKHTPPAREDEPLKRYTFWEVNLKEKFSADLDQFLGRKFLVQAGLRKRP 505
QY 480 -TIGPRKRSAPSATTSSKPAKR 500
    506 FTLGKRRK-ATPTTSTSTTAKR 526
DB 506 FTLGKRRK-ATPTTSTSTTAKR 526

```

Search completed: October 17, 2003, 11:02:03  
 Job time : 72.0888 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2003, 10:47:19 ; Search time 57.1468 Seconds  
(without alignments)  
1280.394 Million cell updates/sec

File: US-08-913-644-4  
Perfect score: 2405  
Sequence: 1 MVSRRARRKRASVTDLTKT.....FLPKRRKRVYFADGFWAA 461

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 45 summaries

Database ::  
1: A.Geneseq.19Jun03.\*  
2: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:\*  
3: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:\*  
4: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:\*  
5: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:\*  
6: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:\*  
7: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:\*  
8: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:\*  
9: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:\*  
10: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:\*  
11: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:\*  
12: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:\*  
13: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:\*  
14: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:\*  
15: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:\*  
16: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:\*  
17: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:\*  
18: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:\*  
19: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:\*  
20: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:\*  
21: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:\*  
23: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:\*  
24: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	2394.5	99.6	462	17	AAW05844	Human papillomavirus
2	2317.5	96.4	462	21	AAW23528	Human papillomavirus
3	2317.5	96.4	462	22	AAW98430	Human papillomavirus
4	1973	82.0	463	22	AAW98433	Human papillomavirus
5	1486.5	61.8	464	22	AAW98455	Human papillomavirus
6	1164.5	48.4	466	22	AAW98437	Human papillomavirus
7	1149.5	47.8	455	22	AAW98416	Human papillomavirus
8	1137.5	47.3	473	23	ABW7480	Human papillomavirus
9	1134	47.2	601	17	AAW05562	Human papillomavirus

10	1133.5	47.1	473	21	AAW05844	Human papillomavirus
11	1131.5	47.0	473	22	AAW05844	Human papillomavirus
12	1131	47.0	459	22	AAW05844	Human papillomavirus
13	1130	47.0	459	22	AAW05844	Human papillomavirus
14	1128.5	46.9	459	22	AAW05844	Human papillomavirus
15	1126	46.8	601	19	AAW05844	Human papillomavirus
16	1114	46.3	467	20	AAW05844	Human papillomavirus
17	1114	46.3	467	20	AAW05844	Human papillomavirus
18	408.5	17.0	469	20	AAW05844	Human papillomavirus
19	400.5	16.7	469	6	AAW05844	Humanised bovine p
20	384.5	16.0	467	24	AAW05844	Bovine papillomavi
21	292.5	12.2	154	21	AAW05844	Amino acid sequenc
22	193.5	8.1	234	4	AAW05844	Sequence encoded b
23	170.5	7.2	2639	22	AAW05844	Novel human diagno
24	170.5	7.1	745	22	AAW05844	Human testes-deriv
25	170.5	7.1	5179	22	AAW05844	C899P predicted am
26	170.5	7.1	5179	22	AAW05844	Human colon tumour
27	170	7.1	492	22	AAW05844	Human testes-deriv
28	169	7.0	695	22	AAW05844	Human testes-deriv
29	166	6.9	745	22	AAW05844	Human testes-deriv
30	166	6.9	778	22	AAW05844	Human testes-deriv
31	166	6.9	3065	24	AAW05844	Human testes-deriv
32	155	6.4	540	22	AAW05844	Human testes-deriv
33	155	6.4	692	22	AAW05844	Human testes-deriv
34	155	6.4	717	22	AAW05844	Human testes-deriv
35	155	6.4	717	22	AAW05844	Human testes-deriv
36	155	6.4	717	22	AAW05844	Human testes-deriv
37	155	6.4	717	22	AAW05844	Human testes-deriv
38	155	6.4	717	22	AAW05844	Human testes-deriv
39	155	6.4	717	22	AAW05844	Human testes-deriv
40	155	6.4	717	22	AAW05844	Human testes-deriv
41	155	6.4	717	22	AAW05844	Human testes-deriv
42	155	6.4	1795	22	AAW05844	Human testes-deriv
43	152	6.3	1296	23	AAW05844	Drosophila melanog
44	152	6.3	1296	23	AAW05844	Human novel polype
45	151	6.3	907	6	AAW05844	Epstein-Barr virus

## ALIGNMENTS

RESULT 1  
ID AAW05844 standard; Protein; 462 AA.  
AC AAW05844;  
XX  
XX 28-JAN-1997 (first entry)  
DT XX  
DE Human papillomavirus type 18 L2 capsid protein.  
KW HPV-18; L2 gene; capsid protein; vaccine; diagnosis; vector;  
KW antibody; serotyping; cervix carcinoma.  
XX  
OS Human papillomavirus type 18.  
XX  
PN W09629413-A2.  
XX  
PD 26-SEP-1996.  
XX  
PE 18-MAR-1996; 96MO-US03649.  
XX  
PR 22-MAR-1995; 95US-0409122.  
XX  
PR 22-MAR-1995; 95US-0408669.  
XX  
PA (MERI) MERCK & CO INC.  
XX  
PI George HA, Hofmann KJ, Jansen KU, Joyce JG, Nepper MP;  
XX WPI; 1996-443188/44.  
XX N-PSDB; AAT40120.  
XX  
PT DNA encoding human papilloma virus 18, esp. L1 and L2 capsid

proteins - and related vectors and antibodies, useful in protective vaccines, for serotyping HPV infections and as therapeutic agents

Claim 2; Fig 3; 46pp; English.

The L2 minor capsid protein (AA05844) of human papillomavirus type 18 (HPV18) has a mol.wt. of 55-60 kDa and is thought to be internal to the L1 major capsid protein (AA05843) in the viral capsomere. Its amino acid sequence was deduced from a genomic clone (AAT40120) isolated from human cervical carcinoma SW756 cells. Recombinant L2 and L1 proteins can be produced in transformed host cells and used in vaccines for protection against HPV18; this HPV type is associated with invasive carcinomas of the cervix, vagina, vulva and anal canal. Virus-like particles composed of recombinant L1+L2 can be produced in yeast for use as vaccines.

Sequence 462 AA;

Query Match 99.6%; Score 2394.5; DB 17; Length 462;  
Best Local Similarity 99.8%; Pred. No. 2.7e-171;  
Matches 461; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

```

QY 1 MSHRAARRKRASTYDLYKTKQSGTQPSDVNNVEGTTLADKLQNSSLGIFLGAGIG 60
DB 1 MSHRAARRKRASTYDLYKTKQSGTQPSDVNNVEGTTLADKLQNSSLGIFLGAGIG 60
QY 61 TGSSTGRTGYIPLGGSNTVVDVGPTRPPVIEPVGPTDPSIYTLIEDSSVYTSGAR 120
DB 61 TGSSTGRTGYIPLGGSNTVVDVGPTRPPVIEPVGPTDPSIYTLIEDSSVYTSGAR 120
QY 121 TFGTSGFDITSAGTTTPAVLDITPSSVSISTNTNPNFASDPSIIEVPGTEVSGNV 180
DB 121 TFGTSGFDITSAGTTTPAVLDITPSSVSISTNTNPNFASDPSIIEVPGTEVSGNV 180
QY 121 TFGTSGFDITSAGTTTPAVLDITPSSVSISTNTNPNFASDPSIIEVPGTEVSGNV 180
DB 121 TFGTSGFDITSAGTTTPAVLDITPSSVSISTNTNPNFASDPSIIEVPGTEVSGNV 180
QY 181 FVGPTSGTHGEIEIPLOTFASSGTGEEPISTPLPVRRVAGRLYSRAVOQSVANPE 240
DB 181 FVGPTSGTHGEIEIPLOTFASSGTGEEPISTPLPVRRVAGRLYSRAVOQSVANPE 240
QY 241 FLTRPSSLITYDNDNAFEPVDITLFEPRSNVPSDDFMDIIRLHRPALTSRGTFRSRLG 300
DB 241 FLTRPSSLITYDNDNAFEPVDITLFEPRSNVPSDDFMDIIRLHRPALTSRGTFRSRLG 300
QY 301 QRATMFTRSQTQIGARVHFYHDISPAPSEYIELQPLVSATEDNGLFDIYADDIDPAMP 360
DB 301 QRATMFTRSQTQIGARVHFYHDISPAPSEYIELQPLVSATEDNGLFDIYADDIDPAMP 360
QY 361 VPSRPTSSAVSYSPITSSASSYNTVPLTSSMDVPVYTGPDITLP-TSWPVSPT 419
DB 361 VPSRPTSSAVSYSPITSSASSYNTVPLTSSMDVPVYTGPDITLP-TSWPVSPT 419
QY 420 APASTOYIGIHGTHYILMPLYYFIPKKRKRPVFFADGFVAA 461
DB 420 APASTOYIGIHGTHYILMPLYYFIPKKRKRPVFFADGFVAA 461

```

RESULT 2

AA05844 standard; protein: 462 AA.

AA05844;

18-JAN-2001 (first entry)

Human papillomavirus 18 L2 protein SEQ ID NO:4.

Human papillomavirus; HPV16; HPV18; L1 protein; L2 protein; antiviral;

vaccine; immunisation; immune response; infection; diagnosis.

Human papillomavirus.

MO200054730-A2.

21-SEP-2000.

XX 08-MAR-2000; 2000WO-US06017.  
XX 18-MAR-1999; 99US-0125208.  
XX 12-AUG-1999; 99US-0148544.  
XX (HARD) HARVARD COLLEGE.  
XX (HARR) HARRISON S.  
XX (CHEN) CHEN X.  
PI Harrison S, Chen X;  
XX WPI: 2000-628165/60.  
XX  
XX Composition comprising multimer of human papillomavirus L1 protein  
XX useful as vaccines against human papillomavirus infection, as  
XX diagnostic tools for detecting presence of the virus in biological  
XX samples -  
XX  
XX Disclosure; Page 31; 31pp; English.

The present invention describes a composition (I) comprising a multimer of human papillomavirus (HPV) L1 protein (II) which is truncated at its amino terminus. (I) has antiviral activity and can be used in the production of a vaccine. (I) comprising a multimer of HPV L1 protein (a T=1 icosahedral particle comprising a pentamer of L1 protein, one of which comprises HPV L2 protein) when administered induces an immune response against the L1 protein in the humans and is thus useful for immunising the humans against HPV. (I) is thus useful as vaccines against human papillomavirus infection, as diagnostic tools for detecting the presence of HPV in biological samples and as tools for mapping receptor interactions. The present sequence represents an HPV18 L2 protein sequence which is used in the exemplification of the present invention.

Sequence 462 AA;

Query Match 96.4%; Score 2317.5; DB 21; Length 462;  
Best Local Similarity 96.8%; Pred. No. 1.6e-165;  
Matches 447; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

```

QY 1 MSHRAARRKRASTYDLYKTKQSGTQPSDVNNVEGTTLADKLQNSSLGIFLGAGIG 60
DB 1 MSHRAARRKRASTYDLYKTKQSGTQPSDVNNVEGTTLADKLQNSSLGIFLGAGIG 60
QY 61 TGSSTGRTGYIPLGGSNTVVDVGPTRPPVIEPVGPTDPSIYTLIEDSSVYTSGAR 120
DB 61 TGSSTGRTGYIPLGGSNTVVDVGPTRPPVIEPVGPTDPSIYTLIEDSSVYTSGAR 120
QY 121 TFGTSGFDITSAGTTTPAVLDITPSSVSISTNTNPNFASDPSIIEVPGTEVSGNV 180
DB 121 TFGTSGFDITSAGTTTPAVLDITPSSVSISTNTNPNFASDPSIIEVPGTEVSGNV 180
QY 121 TFGTSGFDITSAGTTTPAVLDITPSSVSISTNTNPNFASDPSIIEVPGTEVSGNV 180
DB 121 TFGTSGFDITSAGTTTPAVLDITPSSVSISTNTNPNFASDPSIIEVPGTEVSGNV 180
QY 181 FVGPTSGTHGEIEIPLOTFASSGTGEEPISTPLPVRRVAGRLYSRAVOQSVANPE 240
DB 181 FVGPTSGTHGEIEIPLOTFASSGTGEEPISTPLPVRRVAGRLYSRAVOQSVANPE 240
QY 241 FLTRPSSLITYDNDNAFEPVDITLFEPRSNVPSDDFMDIIRLHRPALTSRGTFRSRLG 300
DB 241 FLTRPSSLITYDNDNAFEPVDITLFEPRSNVPSDDFMDIIRLHRPALTSRGTFRSRLG 300
QY 301 QRATMFTRSQTQIGARVHFYHDISPAPSEYIELQPLVSATEDNGLFDIYADDIDPAMP 360
DB 301 QRATMFTRSQTQIGARVHFYHDISPAPSEYIELQPLVSATEDNGLFDIYADDIDPAMP 360
QY 361 VPSRPTSSAVSYSPITSSASSYNTVPLTSSMDVPVYTGPDITLP-TSWPVSPT 419
DB 361 VPSRPTSSAVSYSPITSSASSYNTVPLTSSMDVPVYTGPDITLP-TSWPVSPT 419
QY 420 APASTOYIGIHGTHYILMPLYYFIPKKRKRPVFFADGFVAA 461
DB 420 APASTOYIGIHGTHYILMPLYYFIPKKRKRPVFFADGFVAA 461

```



```

RESULT 3
AAB98430
ID   AAB98430 standard; Protein; 462 AA.
XX
AC   AAB98430;
XX
DT   22-AUG-2001 (first entry)
XX
DE   Human papillomavirus protein HPV18 L2.
XX
KW   Human papillomavirus; human leukocyte antigen; HLA; immune response;
KW   HPV; epitope; T cell; identification; vaccine; infection; genital wart;
KW   neoplastic growth; antiviral.
XX
OS   Human papillomavirus.
XX
PN   WO200141799-A1.
XX
PD   14-JUN-2001.
XX
PF   11-DEC-2000; 2000WO-US33549.
XX
PR   10-DEC-1999; 99US-0172705.
XX   15-AUG-2000; 2000US-0641528.
XX
PA   (EPI-M-) EPIIMUNE INC.
XX
PI   Sette A, Sidney J, Southwood S, Chesnut R, Celis E, Grey HM;
XX
DR   WPI, 2001-381497/40.
XX
PT   An isolated human papilloma virus (HPV) epitope, useful in vaccines for
PT   treating HPV infections -
XX
PS   Disclosure: Page 22-23; 756pp; English.
XX
CC   The present invention describes an isolated prepared human papillomavirus
CC   (HPV) epitope (I). (I) has antiviral activity, and can be used in
CC   vaccine production. Peptides and corresponding nucleic acid compositions
CC   from the present invention are useful for stimulating an immune response
CC   to HPV by stimulating the production of CTL or HTL responses,
CC   specifically in the treatment or prophylaxis of HPV infection, in persons
CC   who have not manifested symptoms e.g. genital warts or neoplastic growth.
CC   The peptides can also be used in a tetramer staining assay to assess
CC   peripheral blood mononuclear cells for the presence of antigen-specific
CC   CTLs following exposure to a pathogen or immunogen, and as reagents to
CC   evaluate immune recall responses or evaluate the efficacy of a vaccine.
CC   The vaccine compositions are useful for removing warts or treating HPV
CC   infections. The epitopes for inclusion in an epitope-base vaccine may
CC   be selected from conserved regions of viral or tumour-associated
CC   antigens, which reduces the likelihood of escape mutants, also
CC   immunosuppressive epitopes that may be present in whole antigens can be
CC   avoided with the use of epitope-base vaccines. An additional advantage
CC   is the ability to combine selected epitopes (CTL and HTL) and to modify
CC   the composition of the epitopes achieving enhanced immunogenicity, the
CC   major benefit of the vaccine is that is safe and efficacious. AAB98391
CC   to AAB98477 represent polypeptide sequences used in the exemplification
CC   of the present invention.
XX
SQ   Sequence 462 AA:
XX
Query Match 96.4%; Score 2317.5; DB 22; Length 462;
Best Local Similarity 96.8%; Pred. No. 1.6e-165;
Matches 447; Conservative 5; Mismatches 9; Indels 1; Gaps 1;
XX
QY 1 MVSRAARRKRSYVDLYKTKCKSGTSPDVNKNKVEGTTTLAKTLLIOWSGIFLGIGIG 60
DB 1 MVSRAARRKRSYVDLYKTKCKSGTSPDVNKNKVEGTTTLAKTLLIOWSGIFLGIGIG 60
QY 61 TGSGTGRTGYIPLGSRNTVVDGPTRPVVEIEVGPDPISIVTLIEDSSVYTSGARPP 120
DB 61 TGSGTGRTGYIPLGSRNTVVDGPTRPVVEIEVGPDPISIVTLIEDSSVYTSGARPP 120

```

```

QY 121 TFGTSGFDITAGCTTPAVLIDITPSSISVSISTNFTNPAPSDPSIIEVPTGEGVSGNV 180
DB 121 TFGTSGFDITAGCTTPAVLIDITPSSISVSISTNFTNPAPSDPSIIEVPTGEGVSGNV 180
QY 181 FVGTPTSGTHGEIEIPLQFASSGGEPISSSTPLPTVRVAGPRLXSAQVSAVNE 240
DB 181 FVGTPTSGTHGEIEIPLQFASSGGEPISSSTPLPTVRVAGPRLXSAQVSAVNE 240
QY 241 FLTRPSLLTYNPAPEPVDITLTFEPRNSVNDSPDMDIRLRPALTSRGTVEFSRLG 300
DB 241 FLTRPSLLTYNPAPEPVDITLTFEPRNSVNDSPDMDIRLRPALTSRGTVEFSRLG 300
QY 301 QRAFMFRSGTQIGARVHFYHDIISPAPEXIEIQLPLVASATDNGLFYDADIDPAMP 360
DB 301 QRAFMFRSGTQIGARVHFYHDIISPAPEXIEIQLPLVASATDNGLFYDADIDPAMP 360
QY 361 VPSRPTSSAVSTSPSTSSASSSNVYPLSSNDVRYTTPDITLP-PTSVWPIVSP 419
DB 361 VPSRPTSSAVSTSPSTSSASSSNVYPLSSNDVRYTTPDITLP-PTSVWPIVSP 419
QY 420 APASTQYIGIHGTHYIPLMYFYFIPKRRKRVYFFADGFVAA 461
DB 420 APASTQYIGIHGTHYIPLMYFYFIPKRRKRVYFFADGFVAA 462
XX
RESULT 4
AAB98443
ID   AAB98443 standard; Protein; 463 AA.
XX
AC   AAB98443;
XX
DT   22-AUG-2001 (first entry)
XX
DE   Human papillomavirus protein HPV45 L2.
XX
KW   Human papillomavirus; human leukocyte antigen; HLA; immune response;
KW   HPV; epitope; T cell; identification; vaccine; infection; genital wart;
KW   neoplastic growth; antiviral.
XX
OS   Human papillomavirus.
XX
PN   WO200141799-A1.
XX
PD   14-JUN-2001.
XX
PF   11-DEC-2000; 2000WO-US33549.
XX
PR   10-DEC-1999; 99US-0172705.
XX   15-AUG-2000; 2000US-0641528.
XX
PA   (EPI-M-) EPIIMUNE INC.
XX
PI   Sette A, Sidney J, Southwood S, Chesnut R, Celis E, Grey HM;
XX
DR   WPI, 2001-381497/40.
XX
PT   An isolated human papilloma virus (HPV) epitope, useful in vaccines for
PT   treating HPV infections -
XX
PS   Disclosure: Page 25; 756pp; English.
XX
CC   The present invention describes an isolated prepared human papillomavirus
CC   (HPV) epitope (I). (I) has antiviral activity, and can be used in
CC   vaccine production. Peptides and corresponding nucleic acid compositions
CC   from the present invention are useful for stimulating an immune response
CC   to HPV by stimulating the production of CTL or HTL responses,
CC   specifically in the treatment or prophylaxis of HPV infection, in persons
CC   who have not manifested symptoms e.g. genital warts or neoplastic growth.
CC   The peptides can also be used in a tetramer staining assay to assess
CC   peripheral blood mononuclear cells for the presence of antigen-specific
CC   CTLs following exposure to a pathogen or immunogen, and as reagents to
CC   evaluate immune recall responses or evaluate the efficacy of a vaccine.
CC   The vaccine compositions are useful for removing warts or treating HPV

```

CC infections. The epitopes for inclusion in an epitope-base vaccine may  
CC be selected from conserved regions of viral or tumour-associated  
CC antigens, which reduces the likelihood of escape mutants, also  
CC immunosuppressive epitopes that may be present in whole antigens can be  
CC avoided with the use of epitope-base vaccines. An additional advantage  
CC is the ability to combine selected epitopes (CTL and HTL) and to modify  
CC the composition of the epitopes achieving enhanced immunogenicity, the  
CC major benefit of the vaccine is that is safe and efficacious. AAB98391  
CC to AAB98477 represent polypeptide sequences used in the exemplification  
CC of the present invention.

SO Sequence 463 AA;

Query Match 82.0%; Score 1973; DB 22; Length 463;  
Best Local Similarity 81.7%; Pred. No. 1,1e-139;  
Matches 379; Conservative 33; Mismatches 48; Indels 4; Gaps 3;

QY 1 MSHRARARRASTYDLYKTKCKOSGCPSPDVNVEGTTADKLQWSSIGIFLGIGIG 60  
DB 1 MSHRARARRASTYDLYKTKCKOSGCPSPDVNVEGTTADKLQWSSIGIFLGIGIG 60  
QY 61 TSGGTGRTGYIPLGHSNTVDVGPTRPPVIEPVPDPSIVTLIEDSSVTSAGARP 120  
DB 61 TSGGTGRTGYIPLGHSNTVDVGPTRPPVIEPVPDPSIVTLIEDSSVTSAGARP 120  
QY 121 TFGTSGFDITSAGTTTAVLDITPSTSVSISTNTNPAFSDPSIIEVPQTEVSGNV 180  
DB 121 TFGTSGFDITSAGTTTAVLDITPSTSVSISTNTNPAFSDPSIIEVPQTEVSGNV 180  
QY 181 FVGPPTSGTGYEIPLOTFASSGTEPISSTPLPVRNAGRLTSRAVQOVANPE 240  
DB 181 FVGPPTSGTGYEIPLOTFASSGTEPISSTPLPVRNAGRLTSRAVQOVANPE 240  
QY 241 FLTPSSLLTYDNPAPFVDTTLTFEPRSNVPSDFMDIIRLHRPALTSRGTVEFRLG 300  
DB 241 FLTPSSLLTYDNPAPFVDTTLTFEPRSNVPSDFMDIIRLHRPALTSRGTVEFRLG 300  
QY 301 ORATMFTRSQTGARGHYHFDISPIAPSEYIELOLVGATEENGFLPDYADIDPAMP 360  
DB 301 ORATMFTRSQTGARGHYHFDISPIAPSEYIELOLVGATEENGFLPDYADIDPAMP 360  
QY 361 VPSRPTTSSAVSTYSPRISS--ASSYSNVTVPLTSSMDVNVYVGPDTLPP--TSWPIVS 417  
DB 361 VPSRPTTSSAVSTYSPRISS--ASSYSNVTVPLTSSMDVNVYVGPDTLPP--TSWPIVS 417  
QY 418 PLTAPASTQYIGHGTHTYLLPLLYFTPKRRKRVDFEADGFVAA 461  
DB 420 PLTAPASTQYIGHGTHTYLLPLLYFTPKRRKRVDFEADGFVAA 463

## RESULT 5

AAB98455 standard; Protein: 464 AA.

AC AAB98455;  
DT 22-AUG-2001 (first entry)  
DE Human papillomavirus protein HPV56 L2.  
KW Human papillomavirus; human leukocyte antigen; HLA; immune response;  
KM HPV; epitope; T cell; identification; vaccine; infection; genital wart;  
neoplastic growth; antiviral.  
OS Human papillomavirus.  
PN WO20014179-A1.  
PD 14-JUN-2001.  
PF 11-DEC-2000; 2000WO-US33549.  
PR 10-DEC-1999; 99US-01172705.

PR 15-AUG-2000; 2000US-0641528.  
XX (EPIM-) EPIMUNE INC.  
PA Sette A, Sidney J, Southwood S, Chesnut R, Celis E, Grey HM;  
PI WPI: 2001-381497/40.

PT An isolated human papilloma virus (HPV) epitope, useful in vaccines for  
PT treating HPV infections -

PS Disclosure; Page 27; 756pp; English.

CC The present invention describes an isolated prepared human papillomavirus  
CC (HPV) epitope (I). (I) has antiviral activity, and can be used in  
CC vaccine production. Peptides and corresponding nucleic acid compositions  
CC from the present invention are useful for stimulating an immune response  
CC to HPV by stimulating the production of CTL or HTL responses,  
CC specifically in the treatment or prophylaxis of HPV infection, in persons  
CC who have not manifested symptoms e.g. genital warts or neoplastic growth.  
CC The peptides can also be used in a tetramer staining assay to assess  
CC peripheral blood mononuclear cells for the presence of antigen-specific  
CC CTLs following exposure to a pathogen or immunogen, and as reagents to  
CC evaluate immune recall responses or evaluate the efficacy of a vaccine.  
CC The vaccine compositions are useful for removing warts or treating HPV  
CC infections. The epitopes for inclusion in an epitope-base vaccine may  
CC be selected from conserved regions of viral or tumour-associated  
CC antigens, which reduces the likelihood of escape mutants, also  
CC immunosuppressive epitopes that may be present in whole antigens can be  
CC avoided with the use of epitope-base vaccines. An additional advantage  
CC is the ability to combine selected epitopes (CTL and HTL) and to modify  
CC the composition of the epitopes achieving enhanced immunogenicity, the  
CC major benefit of the vaccine is that is safe and efficacious. AAB98391  
CC to AAB98477 represent polypeptide sequences used in the exemplification  
CC of the present invention.

SO Sequence 464 AA;

Query Match 61.8%; Score 1486.5; DB 22; Length 464;  
Best Local Similarity 61.1%; Pred. No. 3.4e-103;  
Matches 284; Conservative 68; Mismatches 108; Indels 5; Gaps 4;

QY 1 MSHRARARRASTYDLYKTKCKOSGCPSPDVNVEGTTADKLQWSSIGIFLGIGIG 60  
DB 1 MSHRARARRASTYDLYKTKCKOSGCPSPDVNVEGTTADKLQWSSIGIFLGIGIG 60  
QY 61 TSGGTGRTGYIPLGHSNTVDVGPTRPPVIEPVPDPSIVTLIEDSSVTSAGARP 120  
DB 61 TSGGTGRTGYIPLGHSNTVDVGPTRPPVIEPVPDPSIVTLIEDSSVTSAGARP 120  
QY 121 TFGTSGFDITSAGTTTAVLDITPSTSVSISTNTNPAFSDPSIIEVPQTEVSGNV 180  
DB 121 TFGTSGFDITSAGTTTAVLDITPSTSVSISTNTNPAFSDPSIIEVPQTEVSGNV 180  
QY 181 FVGPPTSGTGYEIPLOTFASSGTEPISSTPLPVRNAGRLTSRAVQOVANPE 240  
DB 181 FVGPPTSGTGYEIPLOTFASSGTEPISSTPLPVRNAGRLTSRAVQOVANPE 240  
QY 241 FLTPSSLLTYDNPAPFVDTTLTFEPRSNVPSDFMDIIRLHRPALTSRGTVEFRLG 300  
DB 241 FLTPSSLLTYDNPAPFVDTTLTFEPRSNVPSDFMDIIRLHRPALTSRGTVEFRLG 300  
QY 301 ORATMFTRSQTGARGHYHFDISPIAPSEYIELOLVGATEENGFLPDYADIDPAMP 360  
DB 301 ORATMFTRSQTGARGHYHFDISPIAPSEYIELOLVGATEENGFLPDYADIDPAMP 360  
QY 361 VPSRPTTSSAVSTYSPRISS--ASSYSNVTVPLTSSMDVNVYVGPDTLPP--TSWPIVS 417  
DB 361 VPSRPTTSSAVSTYSPRISS--ASSYSNVTVPLTSSMDVNVYVGPDTLPP--TSWPIVS 417  
QY 418 PLTAPASTQYIGHGTHTYLLPLLYFTPKRRKRVDFEADGFVAA 461  
DB 420 PLTAPASTQYIGHGTHTYLLPLLYFTPKRRKRVDFEADGFVAA 464

RESULT 6  
 AAB98437  
 ID AAB98437 standard; Protein; 466 AA.  
 XX  
 AC AAB98437;  
 XX  
 DT 22-AUG-2001 (first entry)  
 XX  
 DE Human papillomavirus protein HPV31 L2.  
 XX  
 KW Human papillomavirus; human leukocyte antigen; HLA; immune response;  
 KW HPV; epitope; T cell; identification; vaccine; infection; genital wart;  
 KW neoplastic growth; antiviral.  
 XX  
 OS Human papillomavirus.  
 XX  
 PN WO200141799-A1.  
 PD 14-JUN-2001.  
 XX  
 PF 11-DEC-2000; 2000WO-US33549.  
 XX  
 PR 10-DEC-1999; 99US-0172705.  
 PR 15-AUG-2000; 2000US-0641528.  
 XX  
 PA (EPIM-) EPIMUNE INC.  
 XX  
 PI Sette A, Sidney J, Southwood S, Chesnut R, Cells E, Grey HM;  
 DR WPI; 2001-381497/40.  
 XX  
 PT An isolated human papilloma virus (HPV) epitope, useful in vaccines for  
 PT treating HPV infections -  
 XX  
 PS Disclosure; page 24; 756pp; English.  
 XX  
 CC The present invention describes an isolated prepared human papillomavirus  
 CC (HPV) epitope (I). (I) has antiviral activity, and can be used in  
 CC vaccine production. Peptides and corresponding nucleic acid compositions  
 CC from the present invention are useful for stimulating an immune response  
 CC to HPV by stimulating the production of CTL or HTL responses,  
 CC specifically in the treatment or prophylaxis of HPV infection, in persons  
 CC who have not manifested symptoms e.g. genital warts or neoplastic growth.  
 CC The peptides can also be used in a tetramer staining assay to assess  
 CC peripheral blood mononuclear cells for the presence of antigen-specific  
 CC CTLs following exposure to a pathogen or immunogen, and as reagents to  
 CC evaluate immune recall responses or evaluate the efficacy of a vaccine.  
 CC The vaccine compositions are useful for removing warts or treating HPV  
 CC infections. The epitopes for inclusion in an epitope-based vaccine may  
 CC be selected from conserved regions of viral or tumour-associated  
 CC antigens, which reduces the likelihood of escape mutants, also  
 CC immunosuppressive epitopes that may be present in whole antigens can be  
 CC avoided with the use of epitope-based vaccines. An additional advantage  
 CC is the ability to combine selected epitopes (CTL and HTL) and to modify  
 CC the composition of the epitopes achieving enhanced immunogenicity, the  
 CC major benefit of the vaccine is that is safe and efficacious. AAB98391  
 CC to AAB98477 represent polypeptide sequences used in the exemplification  
 CC of the present invention.  
 XX  
 Sequence 466 AA;  
 XX

	Query Match	48.4%	Score 1164.5;	DB 22;	Length 466;
	Best Local Similarity	50.6%;	Pred. No. 4.8e-79;		
	Matches 247; Conservative	77;	Mismatches 115;	Indels 49;	Gaps 14;
Oy	1 MYSHRAAR-KRASVTDLYKTKOSGFSPSVNKEGTTLADRLQWSSGIPLGGIGI 59   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :				
Dd	1 MRSKRSTRTKRRASITQLYOTCKAAGTSPSDVIPIRIETTTIADQLIRGSMAVFPGGICAI 60   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :				
Oy	60 GTSGTGGRGYILPGLGRSNTVDVG-PTRPPVYLEPVGPDPDPSIVLTLLIEDSSVTTCAP 118   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :				

```
Db      6I GSGSGGCRGTGYPLPSTRPSVSEASIPRPPVSIDYVGPLDPDSIVLVEESGIVDYGAP 120
Qy      119 RPT--FETIGSGFDITSAGTTTPAVLDITPESSSTVSISTTFNTNPASFSDPSIIIEVPQTGEV 176
          | | | | | : | | | | | : | | | | | : | | | | | : | | | | | : |
Db      121 APTRPHPTTISGFATTAATTADTLADYT-----SVST--HENPTFPDPSYLQPPYAET 172
Qy      177 SGAVFYOTPTSNGHGEELIPTQTFASSCGEEPISTELPTPRVARAGRLTSRAVOQVSV 236
          | | | : : : | | | : : : | | | : : : | | | : : : | | | : : : |
Db      173 SGHLISSSSISHNNEELPMDTFYISTNNKTIESTIPGVRRARAGLYSKATOQVKV 222
Qy      237 ANDEFILRPSSLITYNDNPAFEVY--DTLTFFPRSN--VPDSEFMDIIRLRPALTSRG 292
          : | | : | | | | | | | : : | : | : | | | | | | | | | | | |
Db      233 IDPLEFLSAPQOLTTYENPAVEFYVMNEESIYSFNTSHNIAPBDPDLIALHRPALTSRN 292
Qy      293 TVRFSLRGQRATMFTBSGTQIGARVHFHYDISLIAPSPYEIYLEOPL-----VSATEDNGL 347
          | | | | | : | | | | | | | | | | | | | | | | | | | | | | : : | | |
Db      293 TVRSRLGNQKTARTSGATIGARVHYYYDISINPAGESILEMQGLGASATTTJWDGL 352
Qy      348 FDIYAD-----DIDPAMPVPSRPTTSSANVTSPITSSASSTSNVTPLVTSMD 396
          : | | | | | : : | : | : : | | | | | | | | | | | | | | | |
Db      353 YDIYADDFDTVDTPATHNVSPTSAVGS--TSVASAYVPL-----NTTVLSTGED 400
Qy      397 VPYTYTGEDTL--PLPSWMPVISPRAASTQY-IGIHGTHYLMPLYFIPIKKRRKPYP 455
          : | | : | | | | : | | | | | : | | : | | | | : | | | | | : | | | |
Db      401 IPIFSGDVPIEHAHPQOVPEF--PLAPTPQOVSIFVDGGDFLHSYMYMKRRKRKYVF 458
Qy      454 PADGFVA 461
          | | | | |
Db      459 FTDVSAVA 466
```

RESULT 7  
 ID AAB98416  
 AC AAB98416 standard; Protein; 455 AA.  
 AC AAB98416;  
 DT 22-AUG-2001 (first entry)  
 DE Human papillomavirus protein HPV11 L2.  
 DE Human papillomavirus; human leukocyte antigen; HLA; immune response;  
 KW HPV; epitope; T cell; identification; vaccine; infection; genital wart;  
 KW neoplastic growth; antiviral.  
 OS Human papillomavirus.  
 PM WO200141799-A1.  
 PD 14-JUN-2001.  
 PF 11-DEC-2000; 2000MO-US3549.  
 PR 10-DEC-1999; 99DS-0172705.  
 PR 15-AUG-2000; 2000US-0641528.  
 PA (EPIM-) EPIMKUNE INC.  
 PI Sette A, Sidney J, Southwood S, Chesnut R, Cells E, Grey HM;  
 DR WPI: 2001-381497/40.  
 PT An isolated human papilloma virus (HPV) epitope, useful in vaccines for  
 PT treating HPV infections -  
 PS Disclosure; Page 20; 756pp; English.  
 CC The present invention describes an isolated prepared human papillomavirus  
 CC (HPV) epitope (I). (I) has antiviral activity, and can be used in  
 CC vaccine production. Peptides and corresponding nucleic acid compositions  
 CC from the present invention are useful for stimulating an immune response  
 CC to HPV by stimulating the production of CTL or HTL responses,  
 CC specifically in the treatment or prophylaxis of HPV infection, in persons



```

RESULT 9
AAW03562
ID AAW03562 standard; Protein; 601 AA.
XX
AC AAW03562;
XX
DT 22-JAN-1997 (first entry)
XX
DE Human papillomavirus L2E7 fusion protein.
XX
KW HPV; L2 protein; E7 protein; antigen; vaccine; condyloma; wart;
XX
OS Chimeric human papillomavirus type 6;
XX
OS Chimeric synthetic.
XX
FH Key Location/Qualifiers
FT Peptide 596..601
FT /label= Hexa-histidine_tag
XX
PN WO9626277-A1.
XX
PD 29-AUG-1996.
XX
PF 23-FEB-1996; 96MO-GB00397.
XX
PR 28-JUL-1995; 95GB-0015478.
PR 24-FEB-1995; 95GB-0003786.
PR 08-JUN-1995; 95US-0000034.
XX
PA (CANT-) CANTAB PHARM RES LTD.
XX
PI Carmichael JP, Connor SE, Thompson HSG, Whittle NR;
PI Wilson MJ;
XX
DR N-PSDB; AAT9341.
XX
PT New papilloma virus fusion protein aggregate compns. - comprising
PT the PV L2 and pref. E7 antigens, used in vaccines for prophylaxis or
PT therapy of papilloma virus-associated conditions
XX
PS Disclosure; Fig 1; 45pp; English.
XX
CC An L2E7 fusion protein (AAW03562) comprises a fusion of the L2 and E7
CC proteins of human papillomavirus (HPV). It is encoded by an
CC expression construct (see also AAT9341) obd. by PCR amplification
CC of HPV-6 viral DNA isolated from a clinical isolate of wart tissue.
CC The fusion protein be expressed in eukaryotic or prokaryotic
CC transformed host cells. Other fusions, comprising L2 or L1 with
CC E1, E2, E4, E5 or E6 proteins from other HPV types, can also be
CC produced. The fusion proteins are useful in vaccines and
CC immunogenic compns.
XX
SO Sequence 601 AA;

Query Match 47.2%; Score 1134; DB 17; Length 601;
Best Local Similarity 50.8%; Pred. No. 1.3e-76;
Matches 236; Conservative 75; Mismatches 136; Indels 18; Gaps 10;

```

```

DB 212 APTTSHPIEIPLDIDFVVISSDSGPTSTPVPGTAPR---PVGCLYSRALHQVQYTDPA 268
QY 241 FTRPSSLTITDNPAPFVDTTLTFEPRS--NVPDSDFMDIILHHPALSRGTRFSR 298
DB 269 FLSTPQRLITDNPVYEGEVSVQFSDSIHNNAPDEAFMDIILHHPALSRGTRVYSR 328
QY 299 LQGRATWTRSGTQIGARVHFYHDISPIAPSPXYIELQPLVSAEDNGLEFDIYADDIPA 358
DB 329 IGRGSMWTRSGKHIGARIHIFYDISPIAQAEIEHNPVAAQEDT--FDIAKGFEPD 386
QY 359 MPVPSRPTSSANVTISPTISSAS--YSNVTVPJLTSWDVVPVYTGPDITLPPPTWPIYS 417
DB 387 INPTQHPVYNIISPTVLTSTPTVTOPMGNTPVLSIPNDLFLQSGDPDIFPTAPMGTPRS 446
QY 418 PTPA-PTOYIGIGHYVLYLPLYYFIPKRRKRVFPAFGFYAA 461
DB 447 PVPFALPTGPFVFTTSGGFYHPMYFAKRRKRKIPLEFFSD--VAA 489

RESULT 10
AAB23926
ID AAB23926 standard; protein; 473 AA.
XX
AC AAB23926;
XX
DT 18-JAN-2001 (first entry)
XX
DE Human papillomavirus 16 L2 protein SEQ ID NO:2.
XX
KW Human papillomavirus; HPV16; HPV18; L1 protein; L2 protein; antiviral;
KW vaccine; immunisation; immune response; infection; diagnosis.
XX
OS Human papillomavirus.
XX
PN WO200054730-A2.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000MO-US06017.
XX
PR 18-MAR-1999; 99US-0125208.
PR 12-AUG-1999; 99US-0148544.
XX
PA (HARD) HARVARD COLLEGE.
PA (HARR) HARRISON S.
PA (CHEN/) CHEN X.
XX
PI Harrison S, Chen X;
XX
DR WPI; 2000-628165/60.
XX
PT Composition comprising multimer of human papillomavirus L1 protein
PT useful as vaccines against human papillomavirus infection, as
PT diagnostic tools for detecting presence of the virus in biological
PT samples -
XX
PS Disclosure; Page 30; 31pp; English.
XX
CC The present invention describes a composition (I) comprising a multimer
CC of human papillomavirus (HPV) L1 protein (II) which is truncated at its
CC amino terminus. (I) has antiviral activity and can be used in the
CC production of a vaccine. (I) comprising a multimer of HPV L1 protein
CC (a T=1 icosahedral particle comprising a pentamer of L1 protein, one of
CC which comprises HPV L2 protein) when administered induces an immune
CC response against the L1 protein in the humans and is thus useful for
CC immunising the humans against HPV. (I) is thus useful as vaccines against
CC human papillomavirus infection, as diagnostic tools for detecting the
CC presence of HPV in biological samples and as tools for mapping receptor
CC interactions. The present sequence represents an HPV16 L2 protein
CC sequence which is used in the exemplification of the present invention.
XX
SQ Sequence 473 AA;

```





CC antigens, which reduces the likelihood of escape mutants, also  
 CC immunosuppressive epitopes that may be present in whole antigens can be  
 CC avoided with the use of epitope-based vaccines. An additional advantage  
 CC is the ability to combine selected epitopes (CTL and HTL) and to modify  
 CC the composition of the epitopes achieving enhanced immunogenicity, the  
 CC major benefit of the vaccine is that is safe and efficacious. AAB93391  
 CC of AAB98477 represent polypeptide sequences used in the exemplification  
 CC of the present invention.

XX Sequence 459 AA;

Query Match 47.0%; Score 1130; DB 22; Length 459;  
 Best Local Similarity 50.1%; Pred. No. 1.8e-76;

Matches 235; Conservative 77; Mismatches 139; Indels 18; Gaps 10;

QY 1 MSHRAARAKRASYTDLYKTKOSGTCPSDVNVKVEGTLADKILQMSLGIPTGLGIG 60  
 DB 1 MAHRARRRRKASNTQXCTKCTGCPDVIYKVEHTIADQILKMSLGVFFGGLGIG 60  
 QY 61 TGSCTGRTGTYPLGKSNVVDVGP-TRPPVVEVPGETPDSIVTLIEDSSVYTSAPR 119  
 DB 61 TGSCTGRTGTYPLGTSKAPITSGPMARPPVVEVPASDPSIYSLIEESAILINAGAP 120  
 QY 120 PTFGTSGFDITSGTTPAVLDITPSSVSISTNTNFAFSDPSIIEVPQGEVSGN 179  
 DB 121 IVPAGHGGFTITSSSTTPALDVSIVSH--TTSIFRNVFTEPSTVQPPVEANGH 177  
 QY 180 VFVGTPTSGTHGEIPILOTFASGTCGEPISSTPLP-TVRRVAGPR-LYSRAVOQVSV 236  
 DB 178 ILISAPITTHSPEIPILOTFVIVSSSDGPTSSTPVPTAR--PRGLXSRALHQVQY 234  
 QY 237 ANPELRRPSSLITYDNAPFVDTTLFEPRS--NVPDSFMDIIRLHRPALTSRGTV 294  
 DB 235 TDPAHGLTSPOLITYDNVPEGEVSVQFSHDSIHNADEAFMDIIRLHRPALISRGV 294  
 QY - 295 RFSRLGQRATMFTSSGTOIGARVHYHDISPIASPEYIELOPLVSATEDGLDITVAD 354  
 DB 295 RFSRLGQRATMFTSSGTOIGARVHYHDISPIASPEYIELOPLVSATEDGLDITVAD 354  
 QY 355 IDPAPVPSRPTSSAVTSPTISSASS-YSNVTPLTS-SWDVYVYTGPDITLPTSV 412  
 DB 353 FEEDINPTQHVNTISDYLSTPNTVYQPMGNTVPLSLPNDLFGSGDIFPPTAM 412  
 QY 413 WPIYSPAPASTOYIGHGYIYMLPLYEFPKRRKRPYFFADGFFVA 461  
 DB 413 GTPSPVYALPTGPFVITGSGFLHAPMYFARKRRKRRLPFFSD--VAA 459

RESULT 14

AA878317 standard; Protein: 473 AA.

AA878317;

12-MAR-1996 (first entry)

Human papillomavirus 16 wild-type L2 capsid protein.

XX DNA binding activity; human papillomavirus; HPV16; L2 capsid protein;  
 XX N-terminal mutation; deletion; substitution; virus-like particle; VLP;  
 XX cervical cancer; vaccine.

OS Human papillomavirus 16.

XX Key Location/Qualifiers

FT Misc-difference 50

FT Misc-difference 371 /note= "corresponds to an ACT codon"

XX W09520659-A1.

XX 03-AUG-1995.

XX 31-JAN-1995; 95WO-AU00043.

XX 31-JAN-1994; 94AU-0003588.

XX (CSLC-) CSL LTD.

XX (UYOU) UNIV QUEENSLAND.

XX Frazer I, Zhou J;

XX WPI, 1995-275444/36.

XX N-PSDB; AA095109.

XX Modified papilloma virus L2 protein - which binds a minimal amt. of

XX DNA, used for producing virus-like particles for use in vaccines

XX Disclosure: Fig 8; 35pp; English.

CC The N-terminus of HPV16 L2 capsid protein was identified as being  
 CC important for DNA binding. Certain mutations (deletions or  
 CC substitutions) in the wild-type N-terminal sequence can abolish or  
 CC reduce DNA binding. Mutant HPV16 L2 protein which does not bind, or  
 CC which binds only a minimal amt. of, DNA will be useful for producing  
 CC virus-like particles which only incorporate a minimal amt. of DNA  
 CC and are therefore suitable for use in vaccines.

XX DNA coding for the mutant N-terminal sequences is claimed.

XX Sequence 473 AA;

Query Match 46.9%; Score 1128.5; DB 16; Length 473;  
 Best Local Similarity 49.0%; Pred. No. 2.5e-76;

Matches 236; Conservative 74; Mismatches 131; Indels 41; Gaps -12;

QY 6 AARRKASVPTLYTKKOSGTCPSDVNVKVEGTLADKILQMSLGIPTGLGIGTSGT 65  
 DB 7 AKRRKRASATOLYTKKOSGTCPPDILPKVKGKTIADQLQYSGMGVFFGGLGIGTSGT 66  
 QY 66 GGRGTGTYPLGKSNVVDVGP-TRPPVVEVPGETPDSIVTLIEDSSVYTSAPR--PTF 122  
 DB 67 GGRGTGTYPLGKSNVVDVGP-TRPPVVEVPGETPDSIVTLIEDSSVYTSAPR--PTF 122  
 QY 123 -TGSFPTISAGTTPAVLDITPSSVSISTNTNFAFSDPSIIEVPQGEVSGNF 181  
 DB 127 PPVYSGFSTITSTDTTPALDINNVTYV--TTHNNPTTDPSTVLOPPPAETGGHFT 182  
 QY 182 VGTPTSGTHGEIPILOTFASGTCGEPISSTPLP-TVRRVAGPR-LYSRAVOQVSV 241  
 DB 183 LSSSTISHNTEELPMDTIFVSTPNVYTSSTPLGSRPARLGLSTTQVAVVDPPE 242  
 QY 242 LTRPSSLITYDNPAFE--PVDITLFEPRS--VDSDFMDIIRLHRPALTSRGTV 295  
 DB 243 VTPPTKLITYDNPAVEGIDVNTLYFSNONSINADPDPDLVALHAPALTSRRGIR 302  
 QY 296 FSRGQRATMFTSSGTOIGARVHYHDISPIASPEYIELOPLVSATEDGLDITVAD 343  
 DB 303 YSRIGNQOTLRTSGKSGAKVHYHYDLSIDPAE-IELOITIPSTYTTTSHAASPTSI 361  
 QY 344 DNGLEFDIYADI--DPAMPVPSRPTSSAVTSPTISSASSYSNVTPLTS-SWDVYV 400  
 DB 362 NNGLYDIANDFTDTSTTPVPSVTS--LSGITP-----ANTTIPGGKYNPLV 411  
 QY 401 TGPDTLPTSSVMPVSPAPASTOY-IGHGYIYMLPLYEFPKRRKRPYFFADGFFV 459  
 DB 412 SGPDIPINTDOAPSLPIYVGSQYITLADAGFYIHPASYMKRRKRRLPYFFSDVSL 471  
 QY 460 AA 461  
 DB 472 AA 473

RESULT 15

AAW44291 standard; Protein: 601 AA.



```

XX AA044291;
AC
XX 20-JUL-1998 (first entry)
DT
XX
XX Human papillomavirus L2-E7 fusion protein.
DE
XX
XX HPV-6; L2 protein; E7 protein; immunogen; antigen; vaccine;
KM infection; wart; cervical cancer; therapy.
XX
XX Human papillomavirus type 6.
OS Synthetic.
XX
FH Key
FT Misc-difference 105 Location/Qualifiers
FT /note= "residue is Gln in published sequence"
FT Misc-difference 215 /note= "residue is Val in published sequence"
FT Misc-difference 230 /note= "residue is Val in published sequence"
FT Misc-difference 373 /note= "residue is Asp in published sequence"
FT Misc-difference 381 /note= "residue is Glu in published sequence"
FT Misc-difference 386 /note= "residue is Gly in published sequence"
FT Misc-difference 422 /note= "residue is Leu in published sequence"
FT Misc-difference 544 /note= "residue is Phe in published sequence"
FT
FT
PN WO9804706-A1.
XX
XX 05-FEB-1998.
PD
XX
XX 29-JUL-1996; 96WO-GB01816.
PF
XX
XX 29-JUL-1996; 96WO-GB01816.
PR
XX
XX (CANT-) CANTAB PHARM RES LTD.
PA
XX
XX Carmichael JP, Connor SE, Thompson HSG, Whittle NR;
PI Wilson MJ;
PI
XX
XX MPI: 1998-130694/12.
DR N-PSDB: AAV15138.
XX
XX New papillomavirus polypeptide compositions - comprising antigenic
PT determinants in aggregated form, used for prophylaxis or treatment
PT of chronic HPV infection
XX
XX Disclosure: Fig 1; 46pp; English.
PS
XX
XX This polypeptide comprises a fusion between the L2 and E7 proteins
CC of human papillomavirus (HPV). Genes from the HPV-6 virus were
CC amplified by PCR from a viral DNA sample from wart tissue. The
CC isolated genes were used to construct a gene fusion cassette for
CC the expression of HPV-6 derived protein in a heterologous system
CC see AAV15138). A pelb leader sequence was introduced at the
CC N-terminus of the encoded protein to enhance expression in E. coli
CC cells, and a His tag was introduced at the C-terminus to facilitate
CC purification. Fusion polypeptides and aggregates of polypeptides
CC comprising papillomavirus (PV)-derived antigens can be used as
CC immunogens, e.g. as vaccines for prophylaxis or therapy of a
CC PV-associated condition by generating immune responses, e.g.
CC regression of chronic HPV infections, including genital warts
CC (especially where the products and the infections are based on
CC HPV of types 6 and/or 11) or cervical intra-epithelial neoplasia
CC (especially where the products and the infections are based on HPV
CC types 16 and/or 18) in infected patients.
XX
XX Sequence 601 AA:

```

```

Query Match 46.8%; Score 1126; DB 19; Length 601;
Best Local Similarity 50.5%; Pred. No. 5.1e-76;
Matches 235; Conservative 75; Mismatches 137; Indels 18; Gaps 10;

QY 5 RARRRRASVTDLKXCKCKSCGTCPSDVYKNEVEETTADKLQMSGIGPAGIGITGSG 64
   || ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
DB 35 RARRRRASATQLYQCKLGTGCPDPVIPVEHNTADQLKMGSLGVFFGGIGITGSG 94
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
QY 65 TGGRTGYIPLGGRSNRVVDVGP--TRPPVVEIPGTPDPSIVTLIESSVYTSGAPRPTF 123
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
DB 95 TGGRTGYVPLGTSAKSIITSGPMARSVYVEPAPDPSIVSLIESAIIINAGAPETVPP 154
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
QY 124 GTSGEDITSAGTTTPAVLDTTPSTSVSISTNFTNPAPSDPSIIIVPOTGEVSGNVFVG 183
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
DB 155 AHGGEFTITSETTTPALDVSVTSH---TTSIFRNPVEFTEPSVTQPQPPVEANGHILIS 211
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
QY 184 TPTSGHGYEELPLOFPASGCTGEELISSTPLP--TYRRAVGR--LYSRAYQOVSANPE 240
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
DB 212 APTTSHPIELPDLDFVLISSSDSGPTSSTPVGTAPR--PRVGLYSRALMQVQVTDPA 268
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
QY 241 FLTRPESLITYDNPAFEVDYTLTFEPRS--NVPDSDFMDIIRLHRPALTSRGTVRFSR 298
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
DB 269 FLSTPQLRITYDNPPYEGEDVSVQFSDSHNAPDAFMDIIRLHRPALTSRGLYRISR 328
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
QY 299 LGGRAVMTFRSGTQIGARVHFYHDISPIAPSPYEIELQPLVSATEDNGLEFDIYADDIPA 358
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
DB 329 IGGRGSMHTRSGKHIGARIHFYDISPIAQAAEIEHMLPVAAQEDT--FDIYAKGFED 386
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
QY 359 MPVPSRPTSSNAVSTISPTISSAS--YSNVTVLTSMDVPYTGDTLPPTSWPIYVS 417
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
DB 387 INPTQHPVNIISDTYLTSPNVTQWPGNTVPLSLPNDLFLQSGDPIDFPAPMGTGPS 446
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
QY 418 PTAPA-STOYIGIGHYHYLMPLYFIPKKRKRVFEADGFVAA 461
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||
DB 447 PVTPALPTGVPVETTSGGFYIHPAWYFARRKRRIPLFFSD--VAA 489
   ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : ||||| : |||||

```

Search completed: October 17, 2003, 10:59:09  
Job time : 59.1488 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2003, 11:02:14 ; Search time 43.3378 Seconds  
(without alignments)  
1743.646 Million cell updates/sec

Title: US-08-913-644-4  
Perfect score: 1 MVSRRARRRRASVTDLTKT.....FIPKRRKRVFFADGFAVA 461  
Sequence: 1 MVSRRARRRRASVTDLTKT.....FIPKRRKRVFFADGFAVA 461

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 609560 seqs, 163917102 residues  
Total number of hits satisfying chosen parameters: 609560

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

Published Applications, AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2317.5	96.4	462	11	US-09-991-053-4
2	2317.5	96.4	462	15	US-10-301-260A-4
3	1133.5	47.1	473	11	US-09-991-053-2
4	1133.5	47.1	473	15	US-10-301-260A-2
5	1131	47.0	459	12	US-10-224-999A-3470
6	667	27.7	519	12	US-10-224-999A-3468
7	639.5	26.6	519	12	US-10-224-999A-3471
8	408.5	17.0	469	12	US-10-305-765-6
9	408.5	17.0	469	12	US-10-305-765-6
10	408.5	17.0	469	12	US-10-305-633-6
11	408.5	17.0	469	12	US-10-305-633-6
12	174.5	7.3	528	12	US-09-840-746-20
13	170.5	7.1	5179	9	US-09-922-217-1068
14	170.5	7.1	5179	10	US-09-833-263-1068
15	170.5	7.1	5179	14	US-10-025-380-1068

16	168	7.0	1367	10	US-09-801-368-108	Sequence 108, App
17	154.5	6.4	1443	15	US-10-245-802-18	Sequence 18, Appl
18	151.5	6.3	605	10	US-09-801-368-428	Sequence 428, App
19	149.5	6.2	314	15	US-10-043-487-311	Sequence 311, App
20	149	6.2	1794	12	US-09-965-738-299	Sequence 299, App
21	149	6.2	1799	12	US-09-965-738-149	Sequence 149, App
22	149	6.2	1821	15	US-09-965-738-82	Sequence 82, Appl
23	149	6.2	5877	12	US-10-142-515-11	Sequence 11, Appl
24	149	6.2	5935	15	US-10-243-243A-8	Sequence 8, Appl
25	149	6.2	11721	12	US-09-965-738-162	Sequence 162, App
26	148	6.2	2971	15	US-10-146-473-50	Sequence 50, Appl
27	145.5	6.0	957	12	US-09-840-746-19	Sequence 19, Appl
28	145.5	6.0	1322	10	US-09-801-368-114	Sequence 114, App
29	144.5	6.0	957	9	US-09-922-217-1065	Sequence 1065, Ap
30	144.5	6.0	957	10	US-09-833-263-1065	Sequence 1065, Ap
31	144.5	6.0	957	14	US-10-025-380-1065	Sequence 1065, Ap
32	144	6.0	1537	10	US-09-801-368-104	Sequence 104, App
33	141	5.9	1260	15	US-10-245-802-8	Sequence 8, Appl
34	140.5	5.8	688	9	US-09-864-761-36047	Sequence 36047, A
35	140	5.8	2468	12	US-10-246-330-4	Sequence 4, Appl
36	139.5	5.7	1047	15	US-10-245-802-22	Sequence 22, Appl
37	136.5	5.6	1728	12	US-10-032-585-7641	Sequence 7641, Ap
38	134.5	5.5	941	14	US-10-124-557-14	Sequence 14, Appl
39	131.5	5.5	1022	14	US-10-124-557-84	Sequence 84, Appl
40	131.5	5.5	1038	14	US-10-124-557-74	Sequence 74, Appl
41	131.5	5.5	1049	14	US-10-124-557-58	Sequence 58, Appl
42	131.5	5.5	1056	14	US-10-161-510-10	Sequence 10, Appl
43	131.5	5.5	1140	14	US-10-124-557-104	Sequence 104, Appl
44	131.5	5.5	1270	14	US-10-124-557-44	Sequence 44, Appl
45	131.5	5.5				

#### ALIGNMENTS

RESULT 1  
US-09-991-053-4  
; Sequence 4, Application US/09991053  
; Publication No. US20030003532A1  
; GENERAL INFORMATION:  
; APPLICANT: Harrison, Stephen  
; TITLE OF INVENTION: Vaccine compositions and methods for human papilloma  
; FILE REFERENCE: Harvard/Harrison 12687/1120  
; CURRENT APPLICATION NUMBER: US/0991,053  
; CURRENT FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: 09/520,781  
; PRIOR FILING DATE: 2000-03-18  
; PRIOR APPLICATION NUMBER: 60/125208  
; PRIOR FILING DATE: 1999-03-18  
; PRIOR APPLICATION NUMBER: 60/148544  
; PRIOR FILING DATE: 1999-08-12  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 462  
; TYPE: PRT  
; ORGANISM: Human papillomavirus type 18  
US-09-991-053-4

Query Match 96.4%; Score 2317.5; DB 11; Length 462;  
Best local Similarity 96.8%; Pred. No. 4.2e-182;  
Matches 447; Conservative 5; Mismatches 9; Indels 1; Gaps 1;  
QY 1 MVSRRARRRRASVTDLTKTKKSGRCPSVYVNVKVEGTTLADKIILOMSSIGITLGLGIG 60  
DB 1 MVSRRARRRRASVTDLTKTKKSGRCPSVYVNVKVEGTTLADKIILOMSSIGITLGLGIG 60  
QY TSGGTGRCGYITLGLGRSNTVVDGPTRPVTVIEPGFDPSPITVTLIEDSSVYVTSAPRP 120  
DB 61 TSGGTGRCGYITLGLGRSNTVVDGPTRPVTVIEPGFDPSPITVTLIEDSSVYVTSAPRP 120

```

QY 121 TFGTSGFDITAGTTPAVLIDITPSSVSISTTNFNPAFSDPSIIEVQGTGVCNV 180
| | | | |
DB 121 TFGTSGFDITAGTTPAVLIDITPSSVSISTTNFNPAFSDPSIIEVQGTGVCNV 180
QY 181 FVGTPTSGTHGEIEIPLQTFASSGTGEPISTPLPYRRVAGPRLYSRAVOVSANPE 240
| | | | |
DB 181 FVGTPTSGTHGEIEIPLQTFASSGTGEPISTPLPYRRVAGPRLYSRAVOVSANPE 240
QY 241 FLTRPSSLIITYDNPAFEVDITLTFEPRSNVPSDPMIIRLHRPALTSRRGTVFRSLG 300
| | | | |
DB 241 FLTRPSSLIITYDNPAFEVDITLTFEPRSNVPSDPMIIRLHRPALTSRRGTVFRSLG 300
QY 301 ORATMFRTSGTQIGARVHFYHDISPIAPSPYEIQLQPVASATENDNGLFDIADMDPAVP 360
| | | | |
DB 301 ORATMFRTSGTQIGARVHFYHDISPIAPSPYEIQLQPVASATENDNGLFDIADMDPAVP 360
QY 361 VPSRPTSSAVSTSPRTSSASSYSNVTPLTSSMDVVPVYTGPDITLPL-PTSVMPVISP 419
| | | | |
DB 361 VPSRPTSSAVSTSPRTSSASSYSNVTPLTSSMDVVPVYTGPDITLPL-PTSVMPVISP 419
QY 420 APASTOYIGHGTHTYVLMPLTYFIPKKRRKRVYFFADGFVA 461
| | | | |
DB 420 APASTOYIGHGTHTYVLMPLTYFIPKKRRKRVYFFADGFVA 461
QY 421 APASTOYIGHGTHTYVLMPLTYFIPKKRRKRVYFFADGFVA 462
| | | | |
DB 421 APASTOYIGHGTHTYVLMPLTYFIPKKRRKRVYFFADGFVA 462

```

```

RESULT 2
US-10-301-260A-4
; Sequence 4, Application US/10301260A
; Publication No. US20030118609A1
; GENERAL INFORMATION:
; APPLICANT: Harrison, Stephen
; APPLICANT: Chen, Xiaojiang
; TITLE OF INVENTION: Vaccine compositions and methods for human papilloma
; TITLE OF INVENTION: Virus
; FILE REFERENCE: Harvard/Harrison 12687/1123
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: 09/520,822
; PRIOR FILING DATE: 2000-03-18
; PRIOR APPLICATION NUMBER: 60/125208
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/148544
; PRIOR FILING DATE: 1999-08-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 462
; TYPE: PRT
; ORGANISM: Human papillomavirus type 18
US-10-301-260A-4

```

```

Query Match 96.4%; Score 2317.5; DB 15; Length 462;
Best Local Similarity 96.8%; Pred. No. 4,2e-182;
Matches 447; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 1 MVSRRARRRRKRASTYDLYKTKCKOSGCPDQVNVNVEGTTTLADKLQWSSLGIFLGIGIG 60
| | | | |
DB 1 MVSRRARRRRKRASTYDLYKTKCKOSGCPDQVNVNVEGTTTLADKLQWSSLGIFLGIGIG 60
QY 61 TGSCTGRTGYIPLGGRSNVTVDGPTRPVYIEPVGPTDPSIYTLIEDSSVVTSGAPR 120
| | | | |
DB 61 TGSCTGRTGYIPLGGRSNVTVDGPTRPVYIEPVGPTDPSIYTLIEDSSVVTSGAPR 120
QY 121 TFGTSGFDITAGTTPAVLIDITPSSVSISTTNFNPAFSDPSIIEVQGTGVCNV 180
| | | | |
DB 121 TFGTSGFDITAGTTPAVLIDITPSSVSISTTNFNPAFSDPSIIEVQGTGVCNV 180
QY 181 FVGTPTSGTHGEIEIPLQTFASSGTGEPISTPLPYRRVAGPRLYSRAVOVSANPE 240
| | | | |
DB 181 FVGTPTSGTHGEIEIPLQTFASSGTGEPISTPLPYRRVAGPRLYSRAVOVSANPE 240
QY 241 FLTRPSSLIITYDNPAFEVDITLTFEPRSNVPSDPMIIRLHRPALTSRRGTVFRSLG 300
| | | | |
DB 241 FLTRPSSLIITYDNPAFEVDITLTFEPRSNVPSDPMIIRLHRPALTSRRGTVFRSLG 300

```

```

DB 241 FLTRPSSLIITYDNPAFEVDITLTFEPRSNVPSDPMIIRLHRPALTSRRGTVFRSLG 300
QY 301 ORATMFRTSGTQIGARVHFYHDISPIAPSPYEIQLQPVASATENDNGLFDIADMDPAVP 360
| | | | |
DB 301 ORATMFRTSGTQIGARVHFYHDISPIAPSPYEIQLQPVASATENDNGLFDIADMDPAVP 360
QY 361 VPSRPTSSAVSTSPRTSSASSYSNVTPLTSSMDVVPVYTGPDITLPL-PTSVMPVISP 419
| | | | |
DB 361 VPSRPTSSAVSTSPRTSSASSYSNVTPLTSSMDVVPVYTGPDITLPL-PTSVMPVISP 419
QY 420 APASTOYIGHGTHTYVLMPLTYFIPKKRRKRVYFFADGFVA 461
| | | | |
DB 420 APASTOYIGHGTHTYVLMPLTYFIPKKRRKRVYFFADGFVA 461
QY 421 APASTOYIGHGTHTYVLMPLTYFIPKKRRKRVYFFADGFVA 462
| | | | |
DB 421 APASTOYIGHGTHTYVLMPLTYFIPKKRRKRVYFFADGFVA 462

```

```

RESULT 3
US-09-991-053-2
; Sequence 2, Application US/09991053
; Publication No. US200300332A1
; GENERAL INFORMATION:
; APPLICANT: Harrison, Stephen
; APPLICANT: Chen, Xiaojiang
; TITLE OF INVENTION: Vaccine compositions and methods for human papilloma
; TITLE OF INVENTION: Virus
; FILE REFERENCE: Harvard/Harrison 12687/1120
; CURRENT FILING DATE: 2001-11-21
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-18
; PRIOR APPLICATION NUMBER: 60/125208
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/148544
; PRIOR FILING DATE: 1999-08-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 473
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-09-991-053-2

```

```

Query Match 47.1%; Score 1133.5; DB 11; Length 473;
Best Local Similarity 49.2%; Pred. No. 8.8e-85;
Matches 237; Conservative 73; Mismatches 131; Indels 41; Gaps 12;

QY 6 AARRRRASVTDLYKTKCKOSGCPDQVNVNVEGTTTLADKLQWSSLGIFLGIGIGTSGT 65
| | | | |
DB 6 AARRRRASVTDLYKTKCKOSGCPDQVNVNVEGTTTLADKLQWSSLGIFLGIGIGTSGT 65
QY 7 AKRTRRASVTDLYKTKCKOSGCPDQVNVNVEGTTTLADKLQWSSLGIFLGIGIGTSGT 66
| | | | |
DB 7 AKRTRRASVTDLYKTKCKOSGCPDQVNVNVEGTTTLADKLQWSSLGIFLGIGIGTSGT 66
QY 66 GGRGTGYIPLGGRSNVTVDGPTRPVYIEPVGPTDPSIYTLIEDSSVVTSGAPR--PTF 122
| | | | |
DB 66 GGRGTGYIPLGGRSNVTVDGPTRPVYIEPVGPTDPSIYTLIEDSSVVTSGAPR--PTF 122
QY 67 GGRGTGYIPLGGRSNVTVDGPTRPVYIEPVGPTDPSIYTLIEDSSVVTSGAPR--PTF 126
| | | | |
DB 67 GGRGTGYIPLGGRSNVTVDGPTRPVYIEPVGPTDPSIYTLIEDSSVVTSGAPR--PTF 126
QY 123 -TGRSGFDITAGTTPAVLIDITPSSVSISTTNFNPAFSDPSIIEVQGTGVCNV 181
| | | | |
DB 123 -TGRSGFDITAGTTPAVLIDITPSSVSISTTNFNPAFSDPSIIEVQGTGVCNV 181
QY 127 PPVDSGSIITSTDTPTALIDINNTVTV---TTHNNPTFTDPSVLAOPPTPAETGHTT 182
| | | | |
DB 127 PPVDSGSIITSTDTPTALIDINNTVTV---TTHNNPTFTDPSVLAOPPTPAETGHTT 182
QY 182 VGTPTSGTHGEIEIPLQTFASSGTGEPISTPLPYRRVAGPRLYSRAVOVSANPE 241
| | | | |
DB 182 VGTPTSGTHGEIEIPLQTFASSGTGEPISTPLPYRRVAGPRLYSRAVOVSANPE 241
QY 183 LSSSTITHNVEEIEIPLQTFASSGTGEPISTPLPYRRVAGPRLYSRAVOVSANPE 242
| | | | |
DB 183 LSSSTITHNVEEIEIPLQTFASSGTGEPISTPLPYRRVAGPRLYSRAVOVSANPE 242
QY 242 FLTRPSSLIITYDNPAFE--PVDITLTFEPRSN---VPSDPMIIRLHRPALTSRRGTVR 295
| | | | |
DB 242 FLTRPSSLIITYDNPAFE--PVDITLTFEPRSN---VPSDPMIIRLHRPALTSRRGTVR 295
QY 243 VTPPTKLIIFYDNPAFEIGIDVNTLYFSSNDNSINIAADPDLIDIALHRPALTSRRGTVR 302
| | | | |
DB 243 VTPPTKLIIFYDNPAFEIGIDVNTLYFSSNDNSINIAADPDLIDIALHRPALTSRRGTVR 302
QY 296 FSRIGQATMFTSGTQIGARVHFYHDISPIAPSPYEIQLQPVASATENDNGLFDIADMDPAVP 361
| | | | |
DB 296 FSRIGQATMFTSGTQIGARVHFYHDISPIAPSPYEIQLQPVASATENDNGLFDIADMDPAVP 361
QY 303 YSRIGQATMFTSGTQIGARVHFYHDISPIAPSPYEIQLQPVASATENDNGLFDIADMDPAVP 361
| | | | |
DB 303 YSRIGQATMFTSGTQIGARVHFYHDISPIAPSPYEIQLQPVASATENDNGLFDIADMDPAVP 361
QY 344 DNGLFQIYADT---DPAMVPSRPTSSAVSTSPRTSSASSYSNVTPLTSSMDVVPV 400
| | | | |
DB 344 DNGLFQIYADT---DPAMVPSRPTSSAVSTSPRTSSASSYSNVTPLTSSMDVVPV 400
QY 362 NGLIYIDYADTITDITSTTTPVPSVST--LSGITP-----ANTTIFPGAGVYINPLV 411
| | | | |
DB 362 NGLIYIDYADTITDITSTTTPVPSVST--LSGITP-----ANTTIFPGAGVYINPLV 411

```

QY 401 TGGDILPPTSWPIVSPAPASTQY-IGIHGTHYMLPLXYIPKRRKRVPEFADGV 459  
 Db 412 SGPDIPNITDQAPSLIPVGPSTYIIADAGDFYLBPSYMLRKRKRRLPYFFSDVSL 471  
 QY 460 AA 461  
 Db 472 AA 473

RESULT 4  
 US-10-301-260A-2  
 ; Sequence 2, Application US/10301260A  
 ; Publication No. US20030118609A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Harrison, Stephen  
 ; APPLICANT: Chen, Xiaojiang  
 ; TITLE OF INVENTION: Vaccine compositions and methods for human papilloma  
 ; TITLE OF INVENTION: virus  
 ; FILE REFERENCE: Harvard/Harrison 12687/1123  
 ; CURRENT APPLICATION NUMBER: US/10/301,260A  
 ; CURRENT FILING DATE: 2002-11-21  
 ; PRIOR APPLICATION NUMBER: 09/520,822  
 ; PRIOR FILING DATE: 2000-03-18  
 ; PRIOR APPLICATION NUMBER: 60/125208  
 ; PRIOR FILING DATE: 1999-03-18  
 ; PRIOR APPLICATION NUMBER: 60/148544  
 ; PRIOR FILING DATE: 1999-08-12  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO: 2  
 ; LENGTH: 473  
 ; TYPE: PRT  
 ; ORGANISM: Human papillomavirus type 16  
 US-10-301-260A-2

Query Match 47.1%; Score 1133.5; DB 15; Length 473;  
 Best Local Similarity 49.2%; Pred. No. 8.8e-85;  
 Matches 237; Conservative 73; Mismatches 131; Indels 41; Gaps 12;

QY 6 AARRKASVYDLYKTKQSGTCDVYKVEGTTLADKILQMSLIGLIGTSGGT 65  
 Db 7 AKRTKASATQLYKTKQSGTCDVYKVEGTTLADKILQMSLIGLIGTSGGT 66  
 QY 66 GGTGTGTPGAGSNTVVD-VGPRPVVIEPVGPPTDSTVLTLEDSSVYTGAPR-PTF 122  
 Db 67 GGTGTGTPGAGSNTVVD-VGPRPVVIEPVGPPTDSTVLTLEDSSVYTGAPR-PTF 126  
 QY 123 -TGTSGFDTSAGTTTAVLDITPSSVSISTNTNTPAFSDPSIIEVQTGEVGNV 181  
 Db 127 PPVVSFSTITSDTTPALDINNVTYV---TTHNNPTFTDPSVLDQPTPAETGGHT 182  
 QY 182 VGTPTSGTGEYERIPLOTFASGSGTGEPLSTPLPTVRVAGPLXSRAYQVSVANPEF 241  
 Db 183 LSSSTISTHNVEIEIPDTFIVSTNPTVTSSTPIPSRPVARGLVSRTOQYKVVDPAF 242  
 QY 242 LTPSSLIYDNPATF--PVDTLTLEPPSN---VPDSDMDIIRLHRPALTSRGTVR 295  
 Db 243 VTPPTLIYDNPATF--PVDTLTLEPPSN---VPDSDMDIIRLHRPALTSRGTVR 302  
 QY 296 FSRLLGQATMFTSRGTOIGARVHYHDISPIASPEYIELOPLVSATE----- 343  
 Db 303 YSRIGNKQTLRTSRGSGISGAKVHYDLSIDPAEE-IELQTTPTSTYTTTSHASPTSI 361  
 QY 344 DNLGFIYADDI---DPAMPVSRPTSSAVSTYPTISSASSVNTVPLTSSMDVYV 400  
 Db 362 NNLGFIYADDI---DPAMPVSRPTSSAVSTYPTISSASSVNTVPLTSSMDVYV 411  
 QY 401 TGGDILPPTSWPIVSPAPASTQY-IGIHGTHYMLPLXYIPKRRKRVPEFADGV 459  
 Db 412 SGPDIPNITDQAPSLIPVGPSTYIIADAGDFYLBPSYMLRKRKRRLPYFFSDVSL 471  
 QY 460 AA 461  
 Db 472 AA 473

Db 472 AA 473

RESULT 5  
 US-10-224-999A-3470  
 ; Sequence 3470, Application US/10224999A  
 ; Publication No. US20030171318A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Myriad Genetics, Inc.  
 ; APPLICANT: Morham, Scott  
 ; APPLICANT: Zavit, Kenton  
 ; APPLICANT: Hobden, Adrian  
 ; TITLE OF INVENTION: Composition and Method for Treating Viral Infection  
 ; FILE REFERENCE: 5004.01  
 ; CURRENT APPLICATION NUMBER: US/10/224,999A  
 ; CURRENT FILING DATE: 2003-03-03  
 ; PRIOR APPLICATION NUMBER: US 60/313,695  
 ; PRIOR FILING DATE: 2001-08-20  
 ; NUMBER OF SEQ ID NOS: 3484  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 3470  
 ; LENGTH: 459  
 ; TYPE: PRT  
 ; ORGANISM: Human papillomavirus type 6b  
 US-10-224-999A-3470

Query Match 47.0%; Score 1131; DB 12; Length 459;  
 Best Local Similarity 50.1%; Pred. No. 1.3e-84;  
 Matches 235; Conservative 77; Mismatches 139; Indels 18; Gaps 10;

QY 1 MSHRAARRKASVYDLYKTKQSGTCDVYKVEGTTLADKILQMSLIGLIGTSGGT 60  
 Db 1 MAHSRAARRKASVYDLYKTKQSGTCDVYKVEGTTLADKILQMSLIGLIGTSGGT 60  
 QY 61 TGGTGRTGYIPLGGSNTVVDVGP--TRPVVIEPVGPPTDSTVLTLEDSSVYTGAPR 119  
 Db 61 TGGTGRTGYIPLGGSNTVVDVGP--TRPVVIEPVGPPTDSTVLTLEDSSVYTGAPR 120  
 QY 120 PPTGTSGFDTSAGTTTAVLDITPSSVSISTNTNTPAFSDPSIIEVQTGEVGNV 179  
 Db 121 TTPAGHGFSTPSTSETTPTALDVSYSHT---TTSFRNPVTEPSVGPPOPVEANG 177  
 QY 180 VFGTPTSGTHGIEIPLOTFASGSGTGEPLSTPLPTVRVAGPLXSRAYQVSVANPEF 236  
 Db 178 ILISAPVYTHPTEEPDLPDFVYSSDSGPTSTPVGTAPR---PRVGLYSALQVOY 234  
 QY 237 ANPEFLRPSLITVDNPATFVDTLTLEPPRS--NPVDSDFMDIIRLHRPALTSRGTVR 294  
 Db 235 TDPALFSTPRLITVDNPATFVDTLTLEPPRS--NPVDSDFMDIIRLHRPALTSRGTVR 294  
 QY 295 RFSRLGQATMFTSRGTOIGARVHYHDISPIASPEYIELOPLVSATEDNGLFDIYAD 354  
 Db 295 RYSRIGRSGMHRSGHIGARHYFDISPIAQAABEIMHPLVAAQDDT--FDIYAES 352  
 QY 355 IDPAMPVSRPTSSAVSTYPTISSASS--YSNTVYVPLTSSMDVYVPTGDIILPPTSW 413  
 Db 353 FEEGINTQAPVYINISPTVLTSTPNVTQPGWGTVPPLSLPNLFLQSGPDIFPPAPMG 412  
 QY 414 PIVSPAPASTQY-IGIHGTHYMLPLXYIPKRRKRVPEFADGVAA 461  
 Db 413 TPFSVTPALPTGPFVITGSGFYLBPSYMLRKRKRRLPYFFSD--VAA 459

RESULT 6  
 US-10-224-999A-3468  
 ; Sequence 3468, Application US/10224999A  
 ; Publication No. US20030171318A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Myriad Genetics, Inc.  
 ; APPLICANT: Morham, Scott  
 ; APPLICANT: Zavit, Kenton  
 ; APPLICANT: Hobden, Adrian  
 ; TITLE OF INVENTION: Composition and Method for Treating Viral Infection

```

: FILE REFERENCE: 5004.01
: CURRENT APPLICATION NUMBER: US/10/224,999A
: CURRENT FILING DATE: 2003-03-03
: PRIOR APPLICATION NUMBER: US 60/313,695
: PRIOR FILING DATE: 2001-08-20
: NUMBER OF SEQ ID NOS: 3484
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 3468
: LENGTH: 519
: TYPE: PRT
: ORGANISM: Human papillomavirus type 23
US-10-224-999A-3468

```

```

Query Match      27.7%; Score 667; DB 12; Length 519;
Best Local Similarity 38.4%; Pred. No. 2.2e-46;
Matches 206; Conservative 55; Mismatches 166; Indels 110; Gaps 26;

```

```

OY 5 RAARRKRSVYDLKTKKOSGTCPSDVYKVEGTTLDKILQMSLGIFFGLGIGTSG 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3 RAQRTKRASVTDIKGCKASGTCPPDVINKVEQNTLADKILQSGVGFGLGIGTSG 62
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 65 TGGRTGYPLPG-----GKSNIVVD-----VGPR--PPVIEVPGTDSIVTLTDS 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 63 TGGATGYPLRPGVGGTPTVVRPAVPEILIGTPELIPVDSIAPIDEARSIVSLTD-- 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 111 SVTSGAPRPTFTGTS-----GPDITSAGTTPAVDI--TPSSTVSISTT 155
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 121 ---SGAADLPSEAEITAEVHPPTVDIGIDTPIVAGRDALILEVDTNPPTRSVTRT 176
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 156 NFNPAFSDPSII--EVPQTGEVS--GNVFGTPTSSTH---GYEIPLOTFASSGTGE 208
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 177 QYDNPSF---QIIESPTITGEASLADHVFEFGSGGHVAGVVEIIDLTPRSYSIEI 233
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 209 PISSTPLETVRRVAGP-----RLYSRAY--QOVSANDEFLTRPSLI--TYD 252
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 234 EEAAPP---RRSTPIERISQEFFRNLRALYNRLTREQOVKNPDLFTTSPKILVRFQD 289
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 253 NPAPFVDTLTFE---PRSNVPSDPMOILRLRLPALT--SRGTAFSRLGQRATFT 307
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 290 NPVDE--EVQIFERDVAEVEPPDRDLIDRLGRPLLTSTEGRLRLSLGQRASLOT 348
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 308 RSGTOIGARVHFYHDISPI--APSEYIEL-----QPLVATEDNGLPDIYA-- 352
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 349 RSGRVSGRVHFYDLSITNEEPIELELGEHSGDASVIEPQSTYIDANLDVEAIQ 408
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 353 DDIDPAMPVPSRP-----TTSASVSTYSPTISSASSYSNVTVP--LTSSWDVP- 398
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 409 DTIDTADYNSADLLDNAIEEFNNSOLVFGTSDRSSASYSIPRESPRETIYVQDIEG 468
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 399 ---VYTGDTLPTSWPIYSPTAPASTOYIGIHGTHYMLPLYFFPKRRKRPY 452
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 469 NOVITPGE--TERFTITFPL--PSAPAVVHTLQKSPDYLP--SLRKRKRKKRY 518
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 7
US-10-224-999A-3471
: Sequence 3471, Application US/10/224,999A
: Publication No. US2003017318A1
: GENERAL INFORMATION:
: APPLICANT: Myriad Genetics, Inc.
: APPLICANT: Morham, Scott
: APPLICANT: Zavit, Kenton
: APPLICANT: Hobden, Adrian
: TITLE OF INVENTION: Composition and Method for Treating Viral Infection
: FILE REFERENCE: 5004.01
: CURRENT APPLICATION NUMBER: US/10/224,999A
: PRIOR FILING DATE: 2003-03-03
: PRIOR APPLICATION NUMBER: US 60/313,695
: PRIOR FILING DATE: 2001-08-20
: NUMBER OF SEQ ID NOS: 3484
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 3471
: LENGTH: 533

```

```

: TYPE: PRT
: ORGANISM: Human papillomavirus type 9
US-10-224-999A-3471

```

```

Query Match      26.6%; Score 639.5; DB 12; Length 533;
Best Local Similarity 35.8%; Pred. No. 4.1e-44;
Matches 200; Conservative 67; Mismatches 147; Indels 145; Gaps 29;

```

```

OY 5 RAARRKRSVYDLKTKKOSGTCPSDVYKVEGTTLDKILQMSLGIFFGLGIGTSG 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3 RAKRTKRASVTDIKGCKASGTCPPDVINKVEHTTLDKILQSGASAGVFGGLGISTGRG 62
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 65 TGGRTGYPLPG-----GKSNIVVD-----VGPR--PPVIEVPGTDSIVT--- 105
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 63 TGGATGYPLGEGGVGVGTPITVRGVIPEILIGTPELIPDLVPRIDIPAPISVGTID 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 106 -----LIEDSVTSGAPRPTFTGSGPDITSAGTTPAVD 142
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 123 STDVLLGELISTEIHHPVVDNAVVDTPVYTER-----RGSS-----ATLE 165
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 143 ITPSSTSV--SISTNFTNPAFSDPSII--EVPQTGEV-----SGNVFGTP--T 186
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 166 VADSPPMRTVARTQYHNPAF---QIIESTPMGESSLADHITVEGSGQLVGPPE 222
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 187 SGTGHEEIPLOTFASS-----GTGERPISSTP-----LPTVRRVAGRLYSRAI--QQ 233
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 223 SYTASSEIEIQEPPSRYSFEIDGTPPRISTPVORAVQSLSLRRA---LYNRRLTEQ 278
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 234 VSNAPFELTRPSLI--TYDNPAFEPVDTLTFEPPSNV---PDSFMDIIRLRPALT 288
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 279 VAVTDPLFLSRSLVQGFQFNPAFDEYVQIFERDLSYVEEPDRQFLVQRLSLRYLT 338
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 289 -SRGTAFSRLGQRATFTRSQTOIGARVHFYHDISPIAPSEYIELQ-----PL 338
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 339 ETPQGYRVSRSLGRRAITRTSGAQVAGVAFYRDLSTI--NTEEPIMQILGEHSGDSTI 397
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 339 VSATEDNGLPDIYADDDPAMPVPSRPTTS-----SANTSTSP---ISSA 381
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 398 VQGVESISVDVNIDEDP--GLEVGROETPSVEDVFNSEDLLEDGEVDESGQLVVGTR 456
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 382 SSYSNVTVPL-----TSSW--DVVYTGPTITLPT--SVPIVSPAPASTOYIGHGT 432
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 457 RSTNLTIVFRRETFRDTSFYQDIQGT---VSTPESQRTDITLFPIDPTVYIHNDT 513
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 433 --HYLMPLYEYIPKRRR 449
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 514 SGDYLHPSLQ--RKRRR 530
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 8
US-10-305-765-6
: Sequence 6, Application US/10305765
: Publication No. US20030175907A1
: GENERAL INFORMATION:
: APPLICANT: FRAZER, Ian Hector
: APPLICANT: ZHOU, Jian
: TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND METHOD FOR SELECTIVELY
: FILE REFERENCE: 210338.0001/US
: CURRENT APPLICATION NUMBER: US/10/305,765
: PRIOR FILING DATE: 1998-07-09
: PRIOR APPLICATION NUMBER: PCT/AU98/00530
: PRIOR FILING DATE: 2000-01-07
: PRIOR APPLICATION NUMBER: US/09/479,645
: PRIOR FILING DATE: 1997-07-09
: PRIOR APPLICATION NUMBER: AU P07765
: PRIOR FILING DATE: 1997-07-09
: PRIOR APPLICATION NUMBER: AU P09467
: PRIOR FILING DATE: 1997-09-11
: NUMBER OF SEQ ID NOS: 219
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 6
: LENGTH: 469

```

```

; TYPE: prt
; ORGANISM: Bovine papillomavirus type 1
; FEATURE:
; OTHER INFORMATION: L2 open reading frame (wild-type)
US-10-305-765-6

```

```

Query Match      17.0%; Score 408.5; DB 12; Length 469;
Best Local Similarity 29.3%; Pred. No. 3.3e-25;
Matches 152; Conservative 59; Mismatches 174; Indels 133; Gaps 24;

```

```

OY 8 RRRASVVDLKYTKCKOSGTCPSDVVNVKVGTTLADKILLOMSGIFLGIGTGS----- 63
DB 6 RKRASADLYLRYCKKAGTCPPDVIRKVGDTIADILKFGGLATILGGLGIGTGSNGV 65
OY 64 GTGRTGYIPL--GGRSNTVVDVGP-----TRPV-----VIEPVGPTDPSI--VT 105
DB 66 AAGGSPRYTPLRTAGSTSSLASIGSRAYTAGTSPISGAGIPLDLTLETGLALRGVYEDIV 125
OY 106 LIEDSSVYTSAGAPRTFTGSGFDITSACTTP-----AVIDTP--SST 148
DB 126 LPEAPALVTPDA---VPADSGIDLSTGDSSTETLITLLEPEGPEDIAYLELOLDRP 181
OY 149 SVSISTNTNTPAFSDPSIIEVPOTGEVSG--NVFVGTPTSGHGEELPLQTFASGNG 206
DB 182 TWQVSNNAVHQSAYHAPLOLQ--SIAETSGLENIFFGSGLDGTGGENIELTYFGS---- 236
OY 207 EEPISSTPLPYRVAVGRPLY--SRAYQVSVANPEFLTRPSSLITYDNPAEPVDTTLT 264
DB 237 --PRISTPRSIASKSGILNMFSCRYYTVPEDEPVF-----SSQTFANPLVEA----- 284
OY 265 PEPRSNVDPDSDFDIIRLRPALTSRGTVRSRGQRATMTRGTOIGARVHEYHDIS 324
DB 285 --EP-----AVLKGPSGRVGLSQYKPDLTLTRSGTEVGPOLHAYVSLIS 326
OY 325 PIAPSEYI-----ELOPLVSATEDNGLFDIYADDIDPA--NP----- 360
DB 327 TTHEDEALPYVDENTOGIAFVPL--HEQNGFEEIEDDSETHRLLPONTSSPVGS 384
OY 361 -----VPSR---PTTSSAVSTY--SPTISSASYSNVTVPLTSSMDVPYVGTDTLPT 410
DB 385 GVARSLIPIREFSATRPVTGVVYGGSPDYSAS-----PVTD---PDSTSPSLVIDT 433
OY 411 SWMPIVPTAPASTGYIGHGHYILMPLYIFPKRK 448
DB 434 TTPPIIIT---IDGHTVLDLYSSNYTLHPSLLRKRRK 467

```

```

RESULT 9
US-10-305-765-8
; Sequence 8, Application US/10305765
; Publication No. US20030175907A1
; GENERAL INFORMATION:
; APPLICANT: FRAZER, Ian Hector
; APPLICANT: ZHOU, Jian
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND METHOD FOR SELECTIVELY
; TITLE OF INVENTION: EXPRESSING A PROTEIN IN A TARGET CELL OR TISSUE
; FILE REFERENCE: 210338.0001/105
; CURRENT FILING DATE: 2002-11-27
; PRIOR FILING DATE: 2002-11-27
; PRIOR FILING DATE: 2002-11-27
; PRIOR FILING DATE: 2000-01-07
; PRIOR FILING DATE: 1998-07-09
; PRIOR FILING DATE: 1997-07-09
; PRIOR FILING DATE: 1997-07-09
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 8
; LENGTH: 469
; TYPE: prt
; ORGANISM: Artificial Sequence

```

```

; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Bovine
; OTHER INFORMATION: Papillomavirus type 1 L2 open reading frame
; OTHER INFORMATION: (humanized)
; FEATURE:
; OTHER INFORMATION: wild-type codons replaced with synonymous codons
; OTHER INFORMATION: used at relatively high frequency by human genes
US-10-305-765-8

```

```

Query Match      17.0%; Score 408.5; DB 12; Length 469;
Best Local Similarity 29.3%; Pred. No. 3.3e-25;
Matches 152; Conservative 59; Mismatches 174; Indels 133; Gaps 24;

```

```

OY 8 RRRASVVDLKYTKCKOSGTCPSDVVNVKVGTTLADKILLOMSGIFLGIGTGS----- 63
DB 6 RKRASADLYLRYCKKAGTCPPDVIRKVGDTIADILKFGGLATILGGLGIGTGSNGV 65
OY 64 GTGRTGYIPL--GGRSNTVVDVGP-----TRPV-----VIEPVGPTDPSI--VT 105
DB 66 AAGGSPRYTPLRTAGSTSSLASIGSRAYTAGTSPISGAGIPLDLTLETGLALRGVYEDIV 125
OY 106 LIEDSSVYTSAGAPRTFTGSGFDITSACTTP-----AVIDTP--SST 148
DB 126 LPEAPALVTPDA---VPADSGIDLSTGDSSTETLITLLEPEGPEDIAYLELOLDRP 181
OY 149 SVSISTNTNTPAFSDPSIIEVPOTGEVSG--NVFVGTPTSGHGEELPLQTFASGNG 206
DB 182 TWQVSNNAVHQSAYHAPLOLQ--SIAETSGLENIFFGSGLDGTGGENIELTYFGS---- 236
OY 207 EEPISSTPLPYRVAVGRPLY--SRAYQVSVANPEFLTRPSSLITYDNPAEPVDTTLT 264
DB 237 --PRISTPRSIASKSGILNMFSCRYYTVPEDEPVF-----SSQTFANPLVEA----- 284
OY 265 PEPRSNVDPDSDFDIIRLRPALTSRGTVRSRGQRATMTRGTOIGARVHEYHDIS 324
DB 285 --EP-----AVLKGPSGRVGLSQYKPDLTLTRSGTEVGPOLHAYVSLIS 326
OY 325 PIAPSEYI-----ELOPLVSATEDNGLFDIYADDIDPA--NP----- 360
DB 327 TTHEDEALPYVDENTOGIAFVPL--HEQNGFEEIEDDSETHRLLPONTSSPVGS 384
OY 361 -----VPSR---PTTSSAVSTY--SPTISSASYSNVTVPLTSSMDVPYVGTDTLPT 410
DB 385 GVARSLIPIREFSATRPVTGVVYGGSPDYSAS-----PVTD---PDSTSPSLVIDT 433
OY 411 SWMPIVPTAPASTGYIGHGHYILMPLYIFPKRK 448
DB 434 TTPPIIIT---IDGHTVLDLYSSNYTLHPSLLRKRRK 467

```

```

RESULT 10
US-10-305-633-6
; Sequence 6, Application US/10305633
; Publication No. US20030182674A1
; GENERAL INFORMATION:
; APPLICANT: FRAZER, Ian Hector
; APPLICANT: ZHOU, Jian
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND METHOD FOR SELECTIVELY
; TITLE OF INVENTION: EXPRESSING A PROTEIN IN A TARGET CELL OR TISSUE
; FILE REFERENCE: 210338.0001/105
; CURRENT FILING DATE: 2002-11-27
; PRIOR FILING DATE: 2002-11-27
; PRIOR FILING DATE: 2000-01-07
; PRIOR FILING DATE: 1998-07-09
; PRIOR FILING DATE: 1997-07-09
; PRIOR FILING DATE: 1997-07-09
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6

```

```
;; LENGTH: 469
;; TYPE: PRT
;; ORGANISM: Bovine papillomavirus type 1
;; FEATURE:
;; OTHER INFORMATION: L2 open reading frame (wild-type)
US-10-305-633-6
```

```
Query Match      17.0% Score 408.5; DB 12; Length 469;
Best Local Similarity 29.3%; Pred. No. 3.3e-25;
Matches 152; Conservative 59; Mismatches 174; Indels 133; Gaps 24;
```

```

QY 8 RRRKASVTDLYKTCQSGTCDSDVNVNVEGTTLADKILQWSLGIPLGIGIGTGS----63
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 6 RVRKASADLYKTCQKAGTCCPDYIRKVEGDTLADKILKKGGLAIYLGIGIGTWSGTV 65
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 64 GTGRTGTYPL--GGRSNTVVDGP-----TRPPV-----VIEPVGPTDPSI---VT 105
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 66 AAGGSPRYTLRTAGSTSSLASIGRAVTAAGRPSIGAGIPDLTLFTLGLALRPGVEDTV 125
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 106 LIEDSVVTSGAPRPTGTGSGPDITSAGTTP-----AVLDTTP--SST 148
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 126 LPEAPALVTPDA---VPADSGLDALSIGDSTETTLTLEPBGPDIAVLELQPLDRP 181
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 149 SVSISTNTNTPAFSDPSIIEVPQTGEVSG--NVFVGTPSGTGHGEIPLQTFASSGTG 206
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 182 TMOYSNVAHQSAAHAPLQLO--SIAETSGLENIFVGSGIGDGTGENIELTYFGS----236
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 207 EEPISSTPLPTVRNAGPRLY--SRAYQOVSANPELTPRSSLITYDNPAFEPVDTTLT 264
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 237 --PRTSPRISJASRGLIMFMSKRYTOVPTEDPEVF---SSQTFANPLXEA-----284
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 265 FEPRSNVPSDSDFMIDILRHRLALTSRGTVAFSLGQATMFTSSGQIGARVIFYDIS 324
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 285 -EP-----AVLKGPSGRGVLGQVYRPDLITRSGTEVGPQLHVRYSLS 326
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 325 PIAPSPXYI-----ELOPLVATEDNGLFDIYADDIDPA---MP-----360
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 327 TIHEDVAIPYVDENTOGIAFVPL--HEBQAGFEIELDGFSETHRLLPONTSSPFGS 384
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 -----VPSR---PTSSASVSTY--SPTISSASSYSNTVPLTSSMDVPVYTGPDITLPT 410
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 385 GVRSLIPTREFSATRPTGVVYTGSPDTYSAS-----PYTD---PDSTSPSLVIDDT 433
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 411 SVMPTVSPAPASFOYIGIHGTHYLLMPLYYFIPIKRR 448
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 434 TTPIIIT---IDGHIVDLSSNTTLAPSLLRKKRRK 467
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

## RESULT 11

```
US-10-305-633-8
; Sequence 8, Application US/10305633
; Publication No. US20030182674A1
; GENERAL INFORMATION:
; APPLICANT: FRAZER, Ian Hector
; APPLICANT: ZHOU, Jian
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND METHOD FOR SELECTIVELY
; FILE OF INVENTION: EXPRESSING A PROTEIN IN A TARGET CELL OR TISSUE
; FILE REFERENCE: 210338.0001/105
; CURRENT APPLICATION NUMBER: US/10/305,633
; PRIOR APPLICATION NUMBER: US/09/479,645
; PRIOR FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: PCT/AU98/00530
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: AU P07765
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: AU P09467
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 469
; TYPE: PRT
```

```
;; ORGANISM: Artificial Sequence
;; FEATURE:
;; OTHER INFORMATION: Description of Artificial Sequence: Bovine
;; OTHER INFORMATION: papillomavirus type 1 L2 open reading frame
;; FEATURE:
;; OTHER INFORMATION: (humanized)
;; OTHER INFORMATION: wild-type codons replaced with synonymous codons
;; OTHER INFORMATION: used at relatively high frequency by human genes
US-10-305-633-8
```

```
Query Match      17.0% Score 408.5; DB 12; Length 469;
Best Local Similarity 29.3%; Pred. No. 3.3e-25;
Matches 152; Conservative 59; Mismatches 174; Indels 133; Gaps 24;
```

```

QY 8 RRRKASVTDLYKTCQSGTCDSDVNVNVEGTTLADKILQWSLGIPLGIGIGTGS----63
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 6 RVRKASADLYKTCQKAGTCCPDYIRKVEGDTLADKILKKGGLAIYLGIGIGTWSGTV 65
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 64 GTGRTGTYPL--GGRSNTVVDGP-----TRPPV-----VIEPVGPTDPSI---VT 105
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 66 AAGGSPRYTLRTAGSTSSLASIGRAVTAAGRPSIGAGIPDLTLFTLGLALRPGVEDTV 125
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 106 LIEDSVVTSGAPRPTGTGSGPDITSAGTTP-----AVLDTTP--SST 148
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 126 LPEAPALVTPDA---VPADSGLDALSIGDSTETTLTLEPBGPDIAVLELQPLDRP 181
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 149 SVSISTNTNTPAFSDPSIIEVPQTGEVSG--NVFVGTPSGTGHGEIPLQTFASSGTG 206
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 182 TMOYSNVAHQSAAHAPLQLO--SIAETSGLENIFVGSGIGDGTGENIELTYFGS----236
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 207 EEPISSTPLPTVRNAGPRLY--SRAYQOVSANPELTPRSSLITYDNPAFEPVDTTLT 264
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 237 --PRTSPRISJASRGLIMFMSKRYTOVPTEDPEVF---SSQTFANPLXEA-----284
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 265 FEPRSNVPSDSDFMIDILRHRLALTSRGTVAFSLGQATMFTSSGQIGARVIFYDIS 324
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 285 -EP-----AVLKGPSGRGVLGQVYRPDLITRSGTEVGPQLHVRYSLS 326
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 325 PIAPSPXYI-----ELOPLVATEDNGLFDIYADDIDPA---MP-----360
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 327 TIHEDVAIPYVDENTOGIAFVPL--HEBQAGFEIELDGFSETHRLLPONTSSPFGS 384
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 361 -----VPSR---PTSSASVSTY--SPTISSASSYSNTVPLTSSMDVPVYTGPDITLPT 410
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 385 GVRSLIPTREFSATRPTGVVYTGSPDTYSAS-----PYTD---PDSTSPSLVIDDT 433
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 411 SVMPTVSPAPASFOYIGIHGTHYLLMPLYYFIPIKRR 448
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 434 TTPIIIT---IDGHIVDLSSNTTLAPSLLRKKRRK 467
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

## RESULT 12

```
US-09-840-746-20
; Sequence 20, Application US/09840746
; Publication No. US20030166501A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; APPLICANT: Honchell, Cynthia D.
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: Mucin-Related Tumor Marker
; FILE REFERENCE: PC-0039 US
; CURRENT APPLICATION NUMBER: US/09/840,746
; PRIOR FILING DATE: 2001-04-23
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PERL Program
; SEQ ID NO 20
; LENGTH: 528
; TYPE: PRT
; ORGANISM: Sus scrofa
; FEATURE:
; NAME/KEY: misc.feature
; OTHER INFORMATION: Genbank ID No. US20030166501A1 g915208
US-09-840-746-20
```





OY	337	-----PLVSATDNGLFDIADIDIPAMV-----PSRPTSSAVTTP-----	377
		:   :   :	
Db	4226	STSPLEST-----TLSTLPPAITEMSTADPSPPTAPTTTSGCHTLLSPPTTTS	4278
OY	378	-----ISSASSANTVLTLSM-----DVPRYGPIITLPFSVMVIVPT-----	419
		:   :   :	
Db	4279	PGTFRTGTTGSSSAPPSTVQTITSMPTPIPLSI-PSI-KRTTGRLRPBSSYLIC	4366
OY	420	-----APASTQYIGIHG-THRYL	436
		:   :   :	
Db	4337	VLANDTYPAEGEYVNGYIGDTCYFV	4361

```

RESULT 15
US-10-025-380-1068
: Sequence 1068, Application US/10025380
: Publication No. US20020182191A1
GENERAL INFORMATION:
APPLICANT: Xu, Jianshun
APPLICANT: Lodes, Michael J.
APPLICANT: Secretist, Heather
APPLICANT: Benson, Darin R.
APPLICANT: Mesgher, Madeline Joy
APPLICANT: Stoik, John A.
APPLICANT: Wang, Tonglong
APPLICANT: Jiang, Yugu
APPLICANT: Smith, Carole L.
APPLICANT: King, Gordon E.
APPLICANT: Wang, Aijun
APPLICANT: Clapper, Jonathan D.
APPLICANT: Skeiky, Yasar A. W.
APPLICANT: Fanger, Gary R.
APPLICANT: Vedvick, Thomas S.
APPLICANT: Carter, Darick
: TITLE OF INVENTION: COMPOUNDS FOR IMMUNOTHERAPY AND DIAGNOSIS
: TITLE OF INVENTION: OF COLON CANCER AND METHODS FOR THEIR USE
FILE REFERENCE: 210121.471C14
CURRENT APPLICATION NUMBER: US/10/025.380
CURRENT FILING DATE: 2001-12-19
NUMBER-OF SEQ ID NOS: 1129
: SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1068
LENGTH: 5179
TYPE: PRT
ORGANISM: Homo sapiens
US-10-025-380-1068

```

Query Match	7.1%;	Score 170.5;	DB 14;	Length 5179;
Best Local Similarity	24.3%;	Pred. No. 0.00034;		
Matches 108; Conservative	42;	Mismatches 163;	Indels 132;	Gaps 23;

[illegible]

```

Db      4226 STSPLNEST-----TLSTLEPALEMTSTAPSPNPAATTTSGCHLSPSPSTTSP 4278
              ||:
              ||:
QY      378 -----ISSASSYNTVTVLISW---DVPTVTGDPITLPPISVMPISPT---- 419
              ||:
Db      4279 PGATPRTGTTGSSAPPTSVQTTTSTAMTPPLPST-PSI-IRTTGLRPPSSVLIC 4336
              ||:
QY      420 -----APASTOYIGIHG-THYYL 436
              ||:
Db      4337 VLNDITYAPGEEYVINGTYDICTIV 4361
              ||:

```

Search completed: October 17, 2003, 11:15:38  
Job time : 45.3378 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2003, 10:59:19 ; Search time 266.694 Seconds  
(without alignments)  
1572.860 Million cell updates/sec

Title: US-08-913-644-4

Perfect score: 2405  
Sequence: 1 MSHRAARRRRASVTDLKTT.....FLPKRRKRVYFADGFWAA 461

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database:

Pending\_Patents\_AA\_Main:\*

- 1: /cgn2\_6/ptodata/1/paa/PCTUS\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/paa/US06\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/paa/US07\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/paa/US08\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/paa/US081\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/paa/US082\_COMB.pep:\*
- 7: /cgn2\_6/ptodata/1/paa/US083\_COMB.pep:\*
- 8: /cgn2\_6/ptodata/1/paa/US084\_COMB.pep:\*
- 9: /cgn2\_6/ptodata/1/paa/US085\_COMB.pep:\*
- 10: /cgn2\_6/ptodata/1/paa/US086\_COMB.pep:\*
- 11: /cgn2\_6/ptodata/1/paa/US087\_COMB.pep:\*
- 12: /cgn2\_6/ptodata/1/paa/US088\_COMB.pep:\*
- 13: /cgn2\_6/ptodata/1/paa/US089\_COMB.pep:\*
- 14: /cgn2\_6/ptodata/1/paa/US090\_COMB.pep:\*
- 15: /cgn2\_6/ptodata/1/paa/US091\_COMB.pep:\*
- 16: /cgn2\_6/ptodata/1/paa/US092\_COMB.pep:\*
- 17: /cgn2\_6/ptodata/1/paa/US093\_COMB.pep:\*
- 18: /cgn2\_6/ptodata/1/paa/US094\_COMB.pep:\*
- 19: /cgn2\_6/ptodata/1/paa/US095\_COMB.pep:\*
- 20: /cgn2\_6/ptodata/1/paa/US096\_COMB.pep:\*
- 21: /cgn2\_6/ptodata/1/paa/US097A\_COMB.pep:\*
- 22: /cgn2\_6/ptodata/1/paa/US097B\_COMB.pep:\*
- 23: /cgn2\_6/ptodata/1/paa/US098\_COMB.pep:\*
- 24: /cgn2\_6/ptodata/1/paa/US099A\_COMB.pep:\*
- 25: /cgn2\_6/ptodata/1/paa/US099B\_COMB.pep:\*
- 26: /cgn2\_6/ptodata/1/paa/US100\_COMB.pep:\*
- 27: /cgn2\_6/ptodata/1/paa/US101\_COMB.pep:\*
- 28: /cgn2\_6/ptodata/1/paa/US102\_COMB.pep:\*
- 29: /cgn2\_6/ptodata/1/paa/US103\_COMB.pep:\*
- 30: /cgn2\_6/ptodata/1/paa/US104\_COMB.pep:\*
- 31: /cgn2\_6/ptodata/1/paa/US106\_COMB.pep:\*
- 32: /cgn2\_6/ptodata/1/paa/US60\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2317.5	96.4	462	1	PCT-US00-33549-41

2	2317.5	96.4	462	25	US-09-991-053-4	Sequence 4, Appl1
3	2317.5	96.4	462	27	US-10-149-136-41	Sequence 41, Appl1
4	2317.5	96.4	462	29	US-10-301-260A-4	Sequence 4, Appl1
5	1973	82.0	463	1	PCT-US00-33549-54	Sequence 54, Appl1
6	1973	82.0	463	27	US-10-149-136-54	Sequence 54, Appl1
7	1486.5	61.8	464	1	PCT-US00-33549-66	Sequence 66, Appl1
8	1486.5	61.8	464	27	US-10-149-136-66	Sequence 66, Appl1
9	1460	60.7	471	1	PCT-US99-18810-9	Sequence 9, Appl1
10	1164.5	48.4	466	1	PCT-US00-33549-48	Sequence 48, Appl1
11	1164.5	48.4	466	27	US-10-149-136-48	Sequence 48, Appl1
12	1149.5	47.8	455	1	PCT-US00-33549-27	Sequence 27, Appl1
13	1149.5	47.8	455	27	US-10-149-136-27	Sequence 27, Appl1
14	1146.5	47.7	488	1	PCT-US03-04473-8	Sequence 8, Appl1
15	1146.5	47.7	488	1	PCT-US03-04473A-8	Sequence 8, Appl1
16	1146.5	47.7	488	1	PCT-US03-04474-8	Sequence 8, Appl1
17	1146.5	47.7	488	1	PCT-US03-04474A-8	Sequence 8, Appl1
18	1146.5	47.7	488	1	PCT-US03-04480-8	Sequence 8, Appl1
19	1146.5	47.7	488	1	PCT-US03-04480A-8	Sequence 8, Appl1
20	1146.5	47.7	488	1	PCT-US03-04516-8	Sequence 8, Appl1
21	1146.5	47.7	488	1	PCT-US03-04516A-8	Sequence 8, Appl1
22	1146.5	47.7	488	29	US-10-367-043-8	Sequence 8, Appl1
23	1146.5	47.7	488	29	US-10-367-043-8	Sequence 8, Appl1
24	1146.5	47.7	488	29	US-10-367-043-8	Sequence 8, Appl1
25	1146.5	47.7	488	29	US-10-367-043-8	Sequence 8, Appl1
26	1133.5	47.1	473	25	US-09-991-053-2	Sequence 2, Appl1
27	1133.5	47.1	473	29	US-10-301-260A-2	Sequence 2, Appl1
28	1131.5	47.0	470	1	PCT-US03-04473-7	Sequence 7, Appl1
29	1131.5	47.0	470	1	PCT-US03-04473A-7	Sequence 7, Appl1
30	1131.5	47.0	470	1	PCT-US03-04474-7	Sequence 7, Appl1
31	1131.5	47.0	470	1	PCT-US03-04474A-7	Sequence 7, Appl1
32	1131.5	47.0	470	1	PCT-US03-04480-7	Sequence 7, Appl1
33	1131.5	47.0	470	1	PCT-US03-04480A-7	Sequence 7, Appl1
34	1131.5	47.0	470	1	PCT-US03-04516-7	Sequence 7, Appl1
35	1131.5	47.0	470	1	PCT-US03-04516A-7	Sequence 7, Appl1
36	1131.5	47.0	470	29	US-10-367-043-7	Sequence 7, Appl1
37	1131.5	47.0	470	29	US-10-367-043-7	Sequence 7, Appl1
38	1131.5	47.0	470	29	US-10-367-043-7	Sequence 7, Appl1
39	1131.5	47.0	470	29	US-10-368-046-7	Sequence 7, Appl1
40	1131.5	47.0	473	1	PCT-US00-33549-34	Sequence 34, Appl1
41	1131.5	47.0	473	27	US-10-149-136-34	Sequence 34, Appl1
42	1131.5	47.0	536	1	PCT-US03-04473-10	Sequence 10, Appl1
43	1131.5	47.0	536	1	PCT-US03-04473A-10	Sequence 10, Appl1
44	1131.5	47.0	536	1	PCT-US03-04474-10	Sequence 10, Appl1
45	1131.5	47.0	536	1	PCT-US03-04474A-10	Sequence 10, Appl1

#### ALIGNMENTS

RESULT 1  
PCT-US00-33549-41  
Sequence 41, Application PC/TUS0033549  
GENERAL INFORMATION:  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Betteban  
APPLICANT: Grey, Howard M.  
TITLE OF INVENTION: Eptimmune Inc.  
TITLE OF INVENTION: Inducing Cellular Immune Responses to Human  
TITLE OF INVENTION: Papillomavirus Using Peptide and Nucleic Acid  
FILE REFERENCE: 018623-016110PC  
CURRENT APPLICATION NUMBER: PCT/US00/33549  
CURRENT FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 60/172,705  
PRIOR FILING DATE: 1999-12-10  
PRIOR APPLICATION NUMBER: US 09/641,528  
PRIOR FILING DATE: 2000-08-15  
NUMBER OF SEQ ID NOS: 86  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 41

```

: LENGTH: 462
: TYPE: PRT
: ORGANISM: Human papillomavirus type 18
PCT-0500-33549-41

Query Match      96.4%; Score 2317.5; DB 1; Length 462;
Best Local Similarity 96.8%; Pred. No. 1.2e-186;
Matches 447; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

OY 1 MSHRAARRKRASTYDLKTKCKQSGTCCPDVNVKVEGTTLADKILQWSSLGIFLGGLGIG 60
DB 1 MSHRAARRKRASTYDLKTKCKQSGTCCPDVNVKVEGTTLADKILQWSSLGIFLGGLGIG 60
OY 61 TGSCTGRTGYIPLGGRSNTVVDGPTRPVPIEVGPTDPSIYTLIEDSSVYTSGARPR 120
DB 61 TGSCTGRTGYIPLGGRSNTVVDGPTRPVPIEVGPTDPSIYTLIEDSSVYTSGARPR 120
OY 121 TFGTSGFDITSAGTTTTPAVLDITPSSSTSVSISTTNFTNPAFSDPSIIEVQTEVAGNV 180
DB 121 TFGTSGFDITSAGTTTTPAVLDITPSSSTSVSISTTNFTNPAFSDPSIIEVQTEVAGNV 180
OY 121 TFGTSGFDITSAGTTTTPAVLDITPSSSTSVSISTTNFTNPAFSDPSIIEVQTEVAGNV 180
DB 121 TFGTSGFDITSAGTTTTPAVLDITPSSSTSVSISTTNFTNPAFSDPSIIEVQTEVAGNV 180
OY 181 FVGPTSGTGHYEELIPLQTFASSSGTGEPISTPLPVRRVAGRLYSRAVQOVSANPE 240
DB 181 FVGPTSGTGHYEELIPLQTFASSSGTGEPISTPLPVRRVAGRLYSRAVQOVSANPE 240
OY 241 FLTRPSSLIYDNPFAFEVDITLTFEPNSNVPDSDFMDIRLHRPALTSRRGTYRFSRLG 300
DB 241 FLTRPSSLIYDNPFAFEVDITLTFEPNSNVPDSDFMDIRLHRPALTSRRGTYRFSRLG 300
OY 301 ORATMFTRSQTQIGARVHFYHDISPAPSEYIELQPLVSATEDNGLFDIYADDIDPAMP 360
DB 301 ORATMFTRSQTQIGARVHFYHDISPAPSEYIELQPLVSATEDNGLFDIYADDIDPAMP 360
OY 361 VPSRPTSSAVSTSPITSSASSTSNVTVPLTSSMDVVPVYTGPDITLP-PTSVMPISPT 419
DB 361 VPSRPTSSAVSTSPITSSASSTSNVTVPLTSSMDVVPVYTGPDITLP-PTSVMPISPT 419
OY 420 APASTQYIGIGHYHLYLMPLYYFIPKKRRKRVYFFADGFVAA 461
DB 420 APASTQYIGIGHYHLYLMPLYYFIPKKRRKRVYFFADGFVAA 462

RESULT 2
US-09-991-053-4
: Sequence 4, Application US/09991053
: GENERAL INFORMATION:
: APPLICANT: Harrison, Stephen
: TITLE OF INVENTION: Vaccine compositions and methods for human papilloma
: FILE REFERENCE: Harvard/Harrison 12687/1120
: CURRENT FILING DATE: 2001-11-21
: PRIOR APPLICATION NUMBER: US/09/520,781
: PRIOR FILING DATE: 2000-03-18
: PRIOR APPLICATION NUMBER: 60/125208
: PRIOR FILING DATE: 1999-03-18
: PRIOR APPLICATION NUMBER: 60/148544
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 4
: LENGTH: 462
: TYPE: PRT
: ORGANISM: Human papillomavirus type 18
US-09-991-053-4

Query Match      96.4%; Score 2317.5; DB 25; Length 462;
Best Local Similarity 96.8%; Pred. No. 1.2e-186;
Matches 447; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

OY 1 MSHRAARRKRASTYDLKTKCKQSGTCCPDVNVKVEGTTLADKILQWSSLGIFLGGLGIG 60
DB 1 MSHRAARRKRASTYDLKTKCKQSGTCCPDVNVKVEGTTLADKILQWSSLGIFLGGLGIG 60

```

```

DB 1 MSHRAARRKRASTYDLKTKCKQSGTCCPDVNVKVEGTTLADKILQWSSLGIFLGGLGIG 60
OY 61 TGSCTGRTGYIPLGGRSNTVVDGPTRPVPIEVGPTDPSIYTLIEDSSVYTSGARPR 120
DB 61 TGSCTGRTGYIPLGGRSNTVVDGPTRPVPIEVGPTDPSIYTLIEDSSVYTSGARPR 120
OY 121 TFGTSGFDITSAGTTTTPAVLDITPSSSTSVSISTTNFTNPAFSDPSIIEVQTEVAGNV 180
DB 121 TFGTSGFDITSAGTTTTPAVLDITPSSSTSVSISTTNFTNPAFSDPSIIEVQTEVAGNV 180
OY 181 FVGPTSGTGHYEELIPLQTFASSSGTGEPISTPLPVRRVAGRLYSRAVQOVSANPE 240
DB 181 FVGPTSGTGHYEELIPLQTFASSSGTGEPISTPLPVRRVAGRLYSRAVQOVSANPE 240
OY 241 FLTRPSSLIYDNPFAFEVDITLTFEPNSNVPDSDFMDIRLHRPALTSRRGTYRFSRLG 300
DB 241 FLTRPSSLIYDNPFAFEVDITLTFEPNSNVPDSDFMDIRLHRPALTSRRGTYRFSRLG 300
OY 301 ORATMFTRSQTQIGARVHFYHDISPAPSEYIELQPLVSATEDNGLFDIYADDIDPAMP 360
DB 301 ORATMFTRSQTQIGARVHFYHDISPAPSEYIELQPLVSATEDNGLFDIYADDIDPAMP 360
OY 361 VPSRPTSSAVSTSPITSSASSTSNVTVPLTSSMDVVPVYTGPDITLP-PTSVMPISPT 419
DB 361 VPSRPTSSAVSTSPITSSASSTSNVTVPLTSSMDVVPVYTGPDITLP-PTSVMPISPT 419
OY 420 APASTQYIGIGHYHLYLMPLYYFIPKKRRKRVYFFADGFVAA 461
DB 420 APASTQYIGIGHYHLYLMPLYYFIPKKRRKRVYFFADGFVAA 462

RESULT 3
US-10-149-136-41
: Sequence 41, Application US/10149136
: GENERAL INFORMATION:
: APPLICANT: Sette, Alessandro
: APPLICANT: Sidney, John
: APPLICANT: Southwood, Scott
: APPLICANT: Chesnut, Robert
: APPLICANT: Cells, Esteban
: APPLICANT: Grey, Howard M.
: TITLE OF INVENTION: Inducing Cellular Immune Responses to Human Papillomavirus
: FILE REFERENCE: 2060.0100002
: CURRENT FILING DATE: 2000-12-11
: PRIOR APPLICATION NUMBER: US/10/149,136
: PRIOR FILING DATE: 2000-12-11
: PRIOR APPLICATION NUMBER: WO PCT/US00/33549
: PRIOR FILING DATE: 2000-08-15
: PRIOR APPLICATION NUMBER: US 60/172,705
: PRIOR FILING DATE: 1999-12-10
: NUMBER OF SEQ ID NOS: 88
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 41
: LENGTH: 462
: TYPE: PRT
: ORGANISM: Human papillomavirus type 18
US-10-149-136-41

Query Match      96.4%; Score 2317.5; DB 27; Length 462;
Best Local Similarity 96.8%; Pred. No. 1.2e-186;
Matches 447; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

OY 1 MSHRAARRKRASTYDLKTKCKQSGTCCPDVNVKVEGTTLADKILQWSSLGIFLGGLGIG 60
DB 1 MSHRAARRKRASTYDLKTKCKQSGTCCPDVNVKVEGTTLADKILQWSSLGIFLGGLGIG 60
OY 61 TGSCTGRTGYIPLGGRSNTVVDGPTRPVPIEVGPTDPSIYTLIEDSSVYTSGARPR 120
DB 61 TGSCTGRTGYIPLGGRSNTVVDGPTRPVPIEVGPTDPSIYTLIEDSSVYTSGARPR 120
OY 121 TFGTSGFDITSAGTTTTPAVLDITPSSSTSVSISTTNFTNPAFSDPSIIEVQTEVAGNV 180
DB 121 TFGTSGFDITSAGTTTTPAVLDITPSSSTSVSISTTNFTNPAFSDPSIIEVQTEVAGNV 180

```

Db 121 TTGSGFDITSGTTPPAVLDTTPSSISVISTFTNPASDPSEIIEPQGEVAGNV 180  
QY 181 FVGTPTSGHGYEIEIPLQTFASSGTEGEEPISSFTPLTVRRVAGPRISRAYQOVSAANE 240  
Db 181 FVGTPTSGHGYEIEIPLQTFASSGTEGEEPISSFTPLTVRRVAGPRISRAYQOVSAANE 240  
QY 241 FLTRPSSLTYNDPAEPVDITLTTPRNSVNDSDMDIIRLHRPALTSRGTVRFSRLG 300  
Db 241 FLTRPSSLTYNDPAEPVDITLTTPRNSVNDSDMDIIRLHRPALTSRGTVRFSRLG 300  
QY 301 ORATMFTRSCTQIGARVHFYHDISPAPSEYTELOPLVSATEDNGLFDIYADDIDPAMP 360  
Db 301 ORATMFTRSCTQIGARVHFYHDISPAPSEYTELOPLVSATEDNGLFDIYADDIDPAMP 360  
QY 361 VPSRPTSSAVSTYSPPTISSASSYSNVTPLTSSMDVPVYTGPDITLP-PTSWPPIVSPT 419  
Db 361 VPSRPTSSAVSTYSPPTISSASSYSNVTPLTSSMDVPVYTGPDITLP-PTSWPPIVSPT 420  
QY 420 APASTOYIGIHGTHYIYMLPLYFIPKKRRKRVPEFFADGFVAA 461  
Db 421 APASTOYIGIHGTHYIYMLPLYFIPKKRRKRVPEFFADGFVAA 462

RESULT 4  
US-10-301-260A-4  
Sequence 4, Application US/10301260A  
GENERAL INFORMATION:  
APPLICANT: Harrison, Stephen  
APPLICANT: Chen, Xiaojiang  
TITLE OF INVENTION: Vaccine compositions and methods for human papilloma  
FILE REFERENCE: Harvard/Harrison 12687/1123  
CURRENT APPLICATION NUMBER: US/10/301,260A  
PRIOR FILING DATE: 2002-11-21  
PRIOR APPLICATION NUMBER: 09/520,822  
PRIOR FILING DATE: 2000-03-18  
PRIOR APPLICATION NUMBER: 60/125208  
PRIOR FILING DATE: 1999-03-18  
PRIOR APPLICATION NUMBER: 60/148544  
PRIOR FILING DATE: 1999-08-12  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 4  
LENGTH: 462  
TYPE: PRT  
ORGANISM: Human papillomavirus type 18  
US-10-301-260A-4

Query Match 96.4%; Score 2317.5; DB 29; Length 462;  
Best Local Similarity 96.8%; Pred. No. 1.2e-186;  
Matches 447; Conservative 5; Mismatches 9; Indels 1; Gaps 1;  
QY 1 MSHRAARRRRASVMDLYKTCQSGCPCSDVYVKKVGGTTLADKILQWSSIGIFLGGLGIG 60  
Db 1 MSHRAARRRRASVMDLYKTCQSGCPCSDVYVKKVGGTTLADKILQWSSIGIFLGGLGIG 60  
QY 61 TSGGTGGRGCIYPLGGRSNTVDVGPTRPPVIEPVPDPSPITVTLIEDSSVYTSAGAPR 120  
Db 61 TSGGTGGRGCIYPLGGRSNTVDVGPTRPPVIEPVPDPSPITVTLIEDSSVYTSAGAPR 120  
QY 121 TTGTSGFDTISAGTTTTPAVLDTTPSSISVISTFTNPASDPSEIIEVPQGEVAGNV 180  
Db 121 TTGTSGFDTISAGTTTTPAVLDTTPSSISVISTFTNPASDPSEIIEVPQGEVAGNV 180  
QY 181 FVGTPTSGHGYEIEIPLQTFASSGTEGEEPISSFTPLTVRRVAGPRISRAYQOVSAANE 240  
Db 181 FVGTPTSGHGYEIEIPLQTFASSGTEGEEPISSFTPLTVRRVAGPRISRAYQOVSAANE 240  
QY 241 FLTRPSSLTYNDPAEPVDITLTTPRNSVNDSDMDIIRLHRPALTSRGTVRFSRLG 300  
Db 241 FLTRPSSLTYNDPAEPVDITLTTPRNSVNDSDMDIIRLHRPALTSRGTVRFSRLG 300

QY 301 ORATMFTRSCTQIGARVHFYHDISPAPSEYTELOPLVSATEDNGLFDIYADDIDPAMP 360  
Db 301 ORATMFTRSCTQIGARVHFYHDISPAPSEYTELOPLVSATEDNGLFDIYADDIDPAMP 360  
QY 361 VPSRPTSSAVSTYSPPTISSASSYSNVTPLTSSMDVPVYTGPDITLP-PTSWPPIVSPT 419  
Db 361 VPSRPTSSAVSTYSPPTISSASSYSNVTPLTSSMDVPVYTGPDITLP-PTSWPPIVSPT 420  
QY 420 APASTOYIGIHGTHYIYMLPLYFIPKKRRKRVPEFFADGFVAA 461  
Db 421 APASTOYIGIHGTHYIYMLPLYFIPKKRRKRVPEFFADGFVAA 462

RESULT 5  
PCT-US00-33549-54  
Sequence 54, Application PC/TUS0033549  
GENERAL INFORMATION:  
APPLICANT: Sette, Alessandro  
APPLICANT: Sidney, John  
APPLICANT: Southwood, Scott  
APPLICANT: Chesnut, Robert  
APPLICANT: Celis, Betteban  
APPLICANT: Grey, Howard M.  
TITLE OF INVENTION: Inducing Cellular Immune Responses to Human  
TITLE OF INVENTION: Papillomavirus Using Peptide and Nucleic Acid  
FILE REFERENCE: 018623-016110PC  
CURRENT APPLICATION NUMBER: PCT/US00/33549  
CURRENT FILING DATE: 2000-12-11  
PRIOR APPLICATION NUMBER: US 60/1172,705  
PRIOR FILING DATE: 1999-12-10  
PRIOR APPLICATION NUMBER: US 09/641,528  
PRIOR FILING DATE: 2000-08-15  
NUMBER OF SEQ ID NOS: 86  
SOFTWARE: Patentln Ver. 2.1  
SEQ ID NO 54  
LENGTH: 463  
TYPE: PRT  
ORGANISM: Human papillomavirus type 45  
PCT-US00-33549-54

Query Match 82.0%; Score 1973; DB 1; Length 463;  
Best Local Similarity 81.7%; Pred. No. 1.7e-157;  
Matches 379; Conservative 33; Mismatches 48; Indels 4; Gaps 3;  
QY 1 MSHRAARRRRASVMDLYKTCQSGCPCSDVYVKKVGGTTLADKILQWSSIGIFLGGLGIG 60  
Db 1 MSHRAARRRRASVMDLYKTCQSGCPCSDVYVKKVGGTTLADKILQWSSIGIFLGGLGIG 60  
QY 61 TSGGTGGRGCIYPLGGRSNTVDVGPTRPPVIEPVPDPSPITVTLIEDSSVYTSAGAPR 120  
Db 61 TSGGTGGRGCIYPLGGRSNTVDVGPTRPPVIEPVPDPSPITVTLIEDSSVYTSAGAPR 120  
QY 121 TTGTSGFDTISAGTTTTPAVLDTTPSSISVISTFTNPASDPSEIIEVPQGEVAGNV 180  
Db 121 TTGTSGFDTISAGTTTTPAVLDTTPSSISVISTFTNPASDPSEIIEVPQGEVAGNV 180  
QY 181 FVGTPTSGHGYEIEIPLQTFASSGTEGEEPISSFTPLTVRRVAGPRISRAYQOVSAANE 240  
Db 181 FVGTPTSGHGYEIEIPLQTFASSGTEGEEPISSFTPLTVRRVAGPRISRAYQOVSAANE 240  
QY 241 FLTRPSSLTYNDPAEPVDITLTTPRNSVNDSDMDIIRLHRPALTSRGTVRFSRLG 300  
Db 241 FLTRPSSLTYNDPAEPVDITLTTPRNSVNDSDMDIIRLHRPALTSRGTVRFSRLG 300  
QY 301 ORATMFTRSCTQIGARVHFYHDISPAPSEYTELOPLVSATEDNGLFDIYADDIDPAMP 360  
Db 301 ORATMFTRSCTQIGARVHFYHDISPAPSEYTELOPLVSATEDNGLFDIYADDIDPAMP 360  
QY 361 VPSRPTSSAVSTYSPPTISSASSYSNVTPLTSSMDVPVYTGPDITLP-PTSWPPIVSPT 419  
Db 361 VPSRPTSSAVSTYSPPTISSASSYSNVTPLTSSMDVPVYTGPDITLP-PTSWPPIVSPT 419

```

QY 418 PTAASQYIGIHGTHYIMPLYIFPKKKRRVYFFADGVAA 461
DB 420 PTNASTTYIGIHGTHYIMPLYIFPKKKRRIPYFFADGVAA 463

RESULT 6
US-10-149-136-54
; Sequence 54, Application US/10149136
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Cells, Esteban
; APPLICANT: Grey, Howard M.
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Human Papillomavirus
; FILE REFERENCE: 2060 0100002
; CURRENT APPLICATION NUMBER: US/10/149,136
; PRIOR FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: WO PCT/US00/33549
; PRIOR FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 09/641,528
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: US 60/172,705
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 54
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Human papillomavirus type 45
US-10-149-136-54

Query Match 82.0%; Score 1973; DB 27; Length 463;
Best Local Similarity 81.7%; Pred. No. 1.7e-157;
Matches 379; Conservative 33; Mismatches 46; Indels 4; Gaps 3;

QY 1 MSHRAARRKRASTYDLYKTKQSGTCDVYNNKVEGTTLADKILQMSLIGIFLGIG 60
DB 1 MSHRAARRKRASTYDLYKTKQSGTCDVYNNKVEGTTLADKILQMSLIGIFLGIG 60
QY 61 TGSCTGRTGYIPLGSGNSNVVDGPPRPVIEPVGPTDPSIYTLIEDSSVYTSAGARP 120
DB 61 TGSCTGRTGYIPLGSGNSNVVDGPPRPVIEPVGPTDPSIYTLIEDSSVYTSAGARP 120
QY 121 TFGTSGFDITSAGTTTAVLDITPSSSTVSISTNTFNPAFSDPSIIEVQTEVSGNV 180
DB 121 TFGTSGFDITSAGTTTAVLDITPSSSTVSISTNTFNPAFSDPSIIEVQTEVSGNV 180
QY 121 TFGTSGFDITSAGTTTAVLDITPSSSTVSISTNTFNPAFSDPSIIEVQTEVSGNV 180
DB 121 TFGTSGFDITSAGTTTAVLDITPSSSTVSISTNTFNPAFSDPSIIEVQTEVSGNV 180
QY 181 FVGTPGTGHEIEIPLQTEFASSGTGEPISTPLPVRRVAGPRLXSRAQOVSANPE 240
DB 181 FVGTPGTGHEIEIPLQTEFASSGTGEPISTPLPVRRVAGPRLXSRAQOVSANPE 240
QY 181 FVGTPGTGHEIEIPLQTEFASSGTGEPISTPLPVRRVAGPRLXSRAQOVSANPE 240
DB 181 FVGTPGTGHEIEIPLQTEFASSGTGEPISTPLPVRRVAGPRLXSRAQOVSANPE 240
QY 241 FLTPSSLIYDNPFAFEVDITLTFEPRSNVPSDPMIDIRLHRPALTSRGTVRSRLG 300
DB 241 FLTPSSLIYDNPFAFEVDITLTFEPRSNVPSDPMIDIRLHRPALTSRGTVRSRLG 300
QY 241 FLTPSSLIYDNPFAFEVDITLTFEPRSNVPSDPMIDIRLHRPALTSRGTVRSRLG 300
DB 241 FLTPSSLIYDNPFAFEVDITLTFEPRSNVPSDPMIDIRLHRPALTSRGTVRSRLG 300
QY 301 QRATMFRSGQIGARVHFYHDISPIAPSPYEIQLQPLVSAETEDNGLFDIYADIDPAMP 360
DB 301 QRATMFRSGQIGARVHFYHDISPIAPSPYEIQLQPLVSAETEDNGLFDIYADIDPAMP 360
QY 301 QRATMFRSGQIGARVHFYHDISPIAPSPYEIQLQPLVSAETEDNGLFDIYADIDPAMP 360
DB 301 QRATMFRSGQIGARVHFYHDISPIAPSPYEIQLQPLVSAETEDNGLFDIYADIDPAMP 360
QY 361 VPSRPTSSAVSYSPFISS--ASSYSNVTVPLTSSMDVYVYTGPDITLPP--TSVPIVS 417
DB 361 VPSRPTSSAVSYSPFISS--ASSYSNVTVPLTSSMDVYVYTGPDITLPP--TSVPIVS 417
QY 418 PTAASQYIGIHGTHYIMPLYIFPKKKRRVYFFADGVAA 461
DB 420 PTNASTTYIGIHGTHYIMPLYIFPKKKRRIPYFFADGVAA 463

```

```

PCT-US00-33549-66
; Sequence 66, Application PC/US0033549
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Cells, Esteban
; APPLICANT: Grey, Howard M.
; APPLICANT: EpiImmune Inc.
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Human
; TITLE OF INVENTION: Papillomavirus Using Peptide and Nucleic Acid
; FILE REFERENCE: 018623-016110PC
; CURRENT APPLICATION NUMBER: PCT/US00/33549
; CURRENT FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: US 09/641,528
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 66
; LENGTH: 464
; TYPE: PRT
; ORGANISM: Human papillomavirus type 56
PCT-US00-33549-66

```

```

Query Match 61.8%; Score 1486.5; DB 1; Length 464;
Best Local Similarity 61.1%; Pred. No. 2.4e-116;
Matches 284; Conservative 68; Mismatches 108; Indels 5; Gaps 4;

QY 1 MSHRAARRKRASTYDLYKTKQSGTCDVYNNKVEGTTLADKILQMSLIGIFLGIG 60
DB 1 MSHRAARRKRASTYDLYKTKQSGTCDVYNNKVEGTTLADKILQMSLIGIFLGIG 60
QY 61 TGSCTGRTGYIPLGSGNSNVVDGPPRPVIEPVGPTDPSIYTLIEDSSVYTSAGARP 120
DB 61 TGSCTGRTGYIPLGSGNSNVVDGPPRPVIEPVGPTDPSIYTLIEDSSVYTSAGARP 120
QY 121 TFGTSGFDITSAGTTTAVLDITPSSSTVSISTNTFNPAFSDPSIIEVQTEVSGNV 180
DB 121 TFGTSGFDITSAGTTTAVLDITPSSSTVSISTNTFNPAFSDPSIIEVQTEVSGNV 180
QY 121 TFGTSGFDITSAGTTTAVLDITPSSSTVSISTNTFNPAFSDPSIIEVQTEVSGNV 180
DB 121 TFGTSGFDITSAGTTTAVLDITPSSSTVSISTNTFNPAFSDPSIIEVQTEVSGNV 180
QY 181 FVGTPGTGHEIEIPLQTEFASSGTGEPISTPLPVRRVAGPRLXSRAQOVSANPE 240
DB 181 FVGTPGTGHEIEIPLQTEFASSGTGEPISTPLPVRRVAGPRLXSRAQOVSANPE 240
QY 181 FVGTPGTGHEIEIPLQTEFASSGTGEPISTPLPVRRVAGPRLXSRAQOVSANPE 240
DB 181 FVGTPGTGHEIEIPLQTEFASSGTGEPISTPLPVRRVAGPRLXSRAQOVSANPE 240
QY 241 FLTPSSLIYDNPFAFEVDITLTFEPRSNVPSDPMIDIRLHRPALTSRGTVRSRLG 300
DB 241 FLTPSSLIYDNPFAFEVDITLTFEPRSNVPSDPMIDIRLHRPALTSRGTVRSRLG 300
QY 241 FLTPSSLIYDNPFAFEVDITLTFEPRSNVPSDPMIDIRLHRPALTSRGTVRSRLG 300
DB 241 FLTPSSLIYDNPFAFEVDITLTFEPRSNVPSDPMIDIRLHRPALTSRGTVRSRLG 300
QY 301 QRATMFRSGQIGARVHFYHDISPIAPSPYEIQLQPLVSAETED--NGLFPIYA--DDIDP 357
DB 301 QRATMFRSGQIGARVHFYHDISPIAPSPYEIQLQPLVSAETED--NGLFPIYA--DDIDP 357
QY 358 AMPVPSRPTSSAVSYSPFISS--ASSYSNVTVPLTSSMDVYVYTGPDITLPP--TSVPIV 416
DB 358 AMPVPSRPTSSAVSYSPFISS--ASSYSNVTVPLTSSMDVYVYTGPDITLPP--TSVPIV 416
QY 417 SPTPASTQYIGIHGTHYIMPLYIFPKKKRRVYFFADGVAA 461
DB 420 POSPYDTHDYIIGSSSFLALMPYFFRRRRRRKRIYFFADGVAA 464

```

```

; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard M.
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Human Papillomavirus
; TITLE OF INVENTION: Using Peptide and Nucleic Acid Compositions
; FILE REFERENCE: 2060.0100002
; CURRENT APPLICATION NUMBER: US/10/149,136
; PCT-US99-18810-9
; PRIOR APPLICATION NUMBER: WO PCT/US00/33549
; PRIOR FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 09/641,528
; PRIOR FILING DATE: 2000-08-15
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; NUMBER OF SEQ ID NOS: 88
; SOFTWARE: PatentIn version 3.1
; LENGTH: 464
; SEQ ID NO 66
; TYPE: PRF
; ORGANISM: Human papillomavirus type 56
US-10-149-136-66

```

```

Query Match      61.8%; Score 1486.5; DB 27; Length 464;
Best Local Similarity 61.1%; Pred. No. 2.4e-116;
Matches 284; Conservative 68; Mismatches 108; Indels 5; Gaps 4;

```

```

QY 1 MSHRARRRRASVTDLYKTKOSGTCPSDVNKKVEGTTLADKILQMSLGIFLGIGIG 60
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MSHRARRRRASVTDLYKTKOSGTCPSDVNKKVEGTTLADKILQMSLGIFLGIGIG 60
QY 61 TSGTGGRGTYIPLGSRNTVVDGPT-RRPVVIEPVGPTDPSIVTLIEDSSVYTSQARRP 120
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 TSGTGGRGTYIPLGSRNTVVDGPT-RRPVVIEPVGPTDPSIVTLIEDSSVYTSQARRP 120
QY 121 TFGTSGFDITSGATTTPAVLDTTPSSSTVSISTNTNPAFSDPSIIEVPGTGEVGNV 180
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 TFGTSGFDITSGATTTPAVLDTTPSSSTVSISTNTNPAFSDPSIIEVPGTGEVGNV 180
QY 121 NFGSGGFETTSSTTPPAVLDTTPSSSTVHVSSTHTNPLFDPPVIEARPGEGVSGNI 180
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 NFGSGGFETTSSTTPPAVLDTTPSSSTVHVSSTHTNPLFDPPVIEARPGEGVSGNI 180
QY 181 FVGTPTSGTHGYEIPLOTFPAGSGGEBEPISSPTPLTVRRVACPRLYSRAVQVSAANPE 240
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 FVGTPTSGTHGYEIPLOTFPAGSGGEBEPISSPTPLTVRRVACPRLYSRAVQVSAANPE 240
QY 181 LSTPSSGHSYIEIPMOTFAVHSGTSTPISSTPIGFRRIARPRLYRKAQOVKVTDP 240
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 LSTPSSGHSYIEIPMOTFAVHSGTSTPISSTPIGFRRIARPRLYRKAQOVKVTDP 240
QY 241 FLTRPSSLTYNDNPAFEPVDTTLTFEPRSNVPSDPMIDRLRLRPALTSRGVRRSRIG 300
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 FLTRPSSLTYNDNPAFEPVDTTLTFEPRSNVPSDPMIDRLRLRPALTSRGVRRSRIG 300
QY 301 QRAATMFTSGTOIGARVHFYHDISPAPSEVTELOPLVSATED-NGLEPDIVA--DDIDP 357
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 301 QRAATMFTSGTOIGARVHFYHDISPAPSEVTELOPLVSATED-NGLEPDIVA--DDIDP 357
QY 301 RKATIGTRGTOIGARVHFYHDISPAPSEVTELOPLVSATED-NGLEPDIVA--DDIDP 359
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 301 RKATIGTRGTOIGARVHFYHDISPAPSEVTELOPLVSATED-NGLEPDIVA--DDIDP 359
QY 358 AMPVPRPTTSSAVSTYSPITSSASTVNTVPLTSSMDVPTVGTGDTILPP-TSVWPIV 416
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 358 AMPVPRPTTSSAVSTYSPITSSASTVNTVPLTSSMDVPTVGTGDTILPP-TSVWPIV 416
QY 417 SPTAPASTOYIGIGHYHYLMPLYFLPKRRKRVPRFFADGFAA 461
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 417 SPTAPASTOYIGIGHYHYLMPLYFLPKRRKRVPRFFADGFAA 461
QY 420 POSPYVTDVYIQQSSPALMPYFFRRRRKRKRIPFFADGDVAA 464
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 420 POSPYVTDVYIQQSSPALMPYFFRRRRKRKRIPFFADGDVAA 464

```

```

RESULT 9
PCT-US99-18810-9
; Sequence 9, Application PC/TUS9918810
; GENERAL INFORMATION:
; APPLICANT: Indiana University
; APPLICANT: Brown, Daron R.
; APPLICANT: McClellan, Tracy L.
; APPLICANT: Bryan, Janine T.
; APPLICANT: Fife, Kenneth H.
; TITLE OF INVENTION: Purified Human Papillomavirus
; FILE REFERENCE: AKTI 9835 PCT
; CURRENT APPLICATION NUMBER: PCT/US99/18810
; PRIOR FILING DATE: 1999-08-17
; EARLIER APPLICATION NUMBER: US 09/135,241
; EARLIER FILING DATE: 1998-08-17

```

```

; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 471
; TYPE: PRF
; ORGANISM: Human Papillomavirus 83
PCT-US99-18810-9

```

```

Query Match      60.7%; Score 1460; DB 1; Length 471;
Best Local Similarity 62.0%; Pred. No. 4.3e-114;
Matches 298; Conservative 61; Mismatches 90; Indels 32; Gaps 10;

```

```

QY 2 VSHRARRRRASVTDLYKTKOSGTCPSDVNKKVEGTTLADKILQMSLGIFLGIGIG 61
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MSHRARRRRASVTDLYKTKOSGTCPSDVNKKVEGTTLADKILQMSLGIFLGIGIG 60
QY 62 GSGTGGRGTYIPLGSRNTVVDGPT-RRPVVIEPVGPTDPSIVTLIEDSSVYTSQARRP 120
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 62 GSGTGGRGTYIPLGSRNTVVDGPT-RRPVVIEPVGPTDPSIVTLIEDSSVYTSQARRP 120
QY 121 TFGTSGFDITSGATTTPAVLDTTPSSSTVSISTNTNPAFSDPSIIEVPGTGEVGNV 180
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 TFGTSGFDITSGATTTPAVLDTTPSSSTVSISTNTNPAFSDPSIIEVPGTGEVGNV 180
QY 181 FVGTPTSGTHGYEIPLOTFPAGSGGEBEPISSPTPLTVRRVACPRLYSRAVQVSA 237
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 FVGTPTSGTHGYEIPLOTFPAGSGGEBEPISSPTPLTVRRVACPRLYSRAVQVSA 237
QY 238 NPEFLRPSLLTYNDNPAFEPVDTTLTFEPRSNVPSDPMIDRLRLRPALTSRGVRRSRIG 295
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 238 NPEFLRPSLLTYNDNPAFEPVDTTLTFEPRSNVPSDPMIDRLRLRPALTSRGVRRSRIG 295
QY 241 DSAPFMSRPAFVYNDNPAFEPVDTTLTFEPRSNVPSDPMIDRLRLRPALTSRGVRRSRIG 299
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 DSAPFMSRPAFVYNDNPAFEPVDTTLTFEPRSNVPSDPMIDRLRLRPALTSRGVRRSRIG 299
QY 296 FSRGGRATMFTSGTOIGARVHFYHDISPAPSEVTELOPLVSATED-NGLEPDIVA--DDIDP 355
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 296 FSRGGRATMFTSGTOIGARVHFYHDISPAPSEVTELOPLVSATED-NGLEPDIVA--DDIDP 355
QY 300 VSRVQGRATMFTSGTOIGARVHFYHDISPAPSEVTELOPLVSATED-NGLEPDIVA--DDIDP 358
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 300 VSRVQGRATMFTSGTOIGARVHFYHDISPAPSEVTELOPLVSATED-NGLEPDIVA--DDIDP 358
QY 356 DPAMPVPRPTTSSAVSTYSPITSSASTVNTVPLTSSMDVPTVGTGDTILPP-TSVWPIV 405
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 356 DPAMPVPRPTTSSAVSTYSPITSSASTVNTVPLTSSMDVPTVGTGDTILPP-TSVWPIV 405
QY 406 -----TLPTSPVPIVSPFAPASTOYIGIGHYHYLMPLYFLPKRRKRVPRFFADGFAA 460
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 406 -----TLPTSPVPIVSPFAPASTOYIGIGHYHYLMPLYFLPKRRKRVPRFFADGFAA 460
QY 419 MPADPSVPVPSNIP-----QSVFIDGTLYLPLNPIFFPKRRKRKRVHVSADGFAA 469
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 419 MPADPSVPVPSNIP-----QSVFIDGTLYLPLNPIFFPKRRKRKRVHVSADGFAA 469
QY 461 A 461
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 461 A 461
   ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

```

RESULT 10
PCT-US00-33549-48
; Sequence 48, Application PC/TUS0033549
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Chesnut, Robert
; APPLICANT: Celis, Esteban
; APPLICANT: Grey, Howard M.
; TITLE OF INVENTION: Inducing Cellular Immune Responses to Human Papillomavirus Using Peptide and Nucleic Acid
; FILE REFERENCE: 018623-016110PC
; CURRENT APPLICATION NUMBER: PCT/US00/33549
; PRIOR FILING DATE: 2000-12-11
; PRIOR APPLICATION NUMBER: US 60/172,705
; PRIOR FILING DATE: 1999-12-10
; PRIOR APPLICATION NUMBER: US 09/641,528
; PRIOR FILING DATE: 2000-08-15
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 48

```

: LENGTH: 466  
 : TYPE: PRT  
 : ORGANISM: Human papillomavirus type 31  
 PCT-US00-33549-48

```

Query Match          48.4%; Score 1164.5; DB 1; Length 466;
Best Local Similarity 50.6%; Pred. No. 4,3e-89;
Matches 247; Conservative 77; Mismatches 115; Indels 49; Gaps 14;

QY      1  MVSHAAARR-KRASTDLDKTKCKGSGTCGSDVYNKVKEGTTLADKLIIQMSSLGIFLGLGI 59
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB       1  MKSKSTRKRKASATOLQYTCKKAAGTCSDDIPKLEHTDTDLIQLRYSGMSMGFFGLGI 60
        ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      60  GTGSGTGRTGYPIPGSGRNTVVYDVG-PTRPPVVEIEVPGPPTSYITLLIEDSVTVSGAP 118
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB       61  GCGSGTGRTGYVPLSTRPSTVSSEASIPIRRPVSIDPVGRLDPSIVSLVBESGVYDGAP 120
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      119 RPT--FTIGSGFDITSAGTTTPAVLDIDPSSTSVSIISTNFTNPASFSPSIIIEVPQTGEV 176
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB       121 APIHPPTTSGDIDATTADITPAALIDVT-----SVST--HENPFTDPDSVLDPPTPAET 172
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      177 SGNVVGCTPTSGTHGEYEELPLOTGFASSGCEEPISSTPTPYRRVNGPRLSRAYOOVS 236
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB       173 SGHLLSSSSISTHYEELPMOTFIYSTNNENITSTPIPGRRPARRLGLYSATGOVKY 232
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      237 ANPEFLTRPSSLIIYDNPAFEVY--DTLTTFEPRS-N-VPDSDFMDIIRLRPALTSRRC 292
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB       233 IDPFLLSAPOKLIYENPAYEFVNABEESLYFENTSHINIPDDDFDLIALHRPALTSRRN 292
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      293 TVRFGRLOARMTFRSGTOIGARVHFYDISPIASPAYTEILOPL-----VSATDNGI 347
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB       293 TVRVSRLENKQTLRIKRSQATIGARHHYYIIDISSINPAGESIEKOPLAGASATTTSTLNDGL 352
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      348 FDIYAD-----DIDPAMPVPSREPTSSANSTVSPRTISSASYSNVTVPLTNSMD 396
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB       353 YDIYADTDFYVDTPATHNVSPSTAVOS---TSASAAYVPT-----NTVPLSTGFD 400
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      397 VPTYNGPDTLT--PPTSWPIYSFPAPASTQY-IGHGHNYHYLMPLYLPKKRKRVYPF 453
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB       401 IPISGPPVPLEHAFTQVPPF--PLAPTTPQVSIFFVGGDFYLHPSYMLKRRKRKVSYP 458
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY      454 FADGEVAA 461
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
DB       459 FTDVSVAA 466
        |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 11
US-10-149-136-48
: Sequence 48, Application US/10149136
: GENERAL INFORMATION:
: APPLICANT: Sette, Alessandro
: APPLICANT: Sidney, John
: APPLICANT: Southwood, Scott
: APPLICANT: Chesnut, Robert
: APPLICANT: Celis, Esteban
: APPLICANT: Grey, Howard M.
: TITLE OF INVENTION: Using Peptide and Nucleic Acid Compositions
: FILE REFERENCE: 2060_0100002
: CURRENT APPLICATION NUMBER: US/10/149,136
: PRIOR FILING DATE: 2000-12-11
: PRIOR APPLICATION NUMBER: WO PCT/US00/33549
: PRIOR FILING DATE: 2000-12-11
: PRIOR APPLICATION NUMBER: US 09/641,528
: PRIOR FILING DATE: 2000-08-15
: PRIOR APPLICATION NUMBER: US 60/172,705
: NUMBER OF SEQ ID NOS: 88
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 48
: LENGTH: 466
: TYPE: PRT
: ORGANISM: Human papillomavirus type 31

```

```

US-10-149-136-48

Query Match          48.4%, Score 1164.5; DB 27; Length 466;
Best Local Similarity 50.6%, Pred. No. 4,3e-89;
Matches 247; Conservative 77; Mismatches 115; Indels 49; Gaps 14;

QY      1 MVSHRAARR-KRAVTDLYKTKCKOSGCPSPVYNKVVESTTLADKLKWSLSIGIFLGIGI 59
Dd      1 MRSKRSTRKKRKRAATOLYLQTKCAAGTCPSIVIKRIETTTTADQLRLKGSKGVFFGGGLGI 60
QY      60 GTGGGTGCRTGYIDLGRSNTVDVG--PTRPPVYLEPVDPDSIVTLIEDSSVTSGAP 118
Dd      61 GSGGTGCGRTGYVPLSTRPSTVSASIRIPRPPSIDPGPLDDPSIVLSVEESGIYDGAP 120
QY      119 RPT--FTGSGEFDITSAGTTTPPAVLDTIPSSTSVISISTNFTNPAFSDDPSIIIEVQTGEV 176
Dd      121 APTHPPTTSGEDFATVADTVTPALLDVY-----SVST--HENPEFDP SVLQPPTPAET 172
QY      177 SGANVFQPTSFGTHGEYERILQTFPAA SSGTGEEPISTFLPTVRRYAGRLSRAYQQYSV 236
Dd      173 SGHLSSSSISSTHNREYELPMDFITLVSNNNENITSSPTIPGVRPARRLGLVSKAQKYV 232
QY      237 ANPEFLTRPSLLITYDNPAEPV--DTLTLEEPSN--VPQSDPDITRLRLRPALTSRG 292
Dd      233 IDPFELAPAKOLLTYENPAFTVAESLYFSNTSHNIAPPDFLDITLHRPALTSRN 292
QY      293 TVRSRSLGQRATMTFRSGTQICARHAFHDISPIAPSEYTELOPL-----VSATEDNGL 347
Dd      293 TVRSRLGNKQKLTLRSGATIGAVHYIIDISSINPAGESILEMGDLGASATTSTLNDGL 352
QY      348 FDIYAD-----DIPAMEVPSAPTSSSAVSITYSPITSSASSTANTVPLTSSWD 396
Dd      353 YDIADMDYFVDTPATHNVSPSTAVQS---TSANVASVPPt-----NTTPVLTSTFD 400
QY      397 VPVYTGPDLTV--PPTSWPPIVSPFPAPASTQY-IGIHSTHYLYMLPYFIKRRKRVYPF 453
Dd      401 IPIRSGDPVPLEHAFTQVFPF--FLAFTTPQVSYFDVGDGFYLPBSTM LKRRKKRYSYF 458
QY      454 PADGFVA A 461
Dd      459 FTDVSAVA 466

RESULT 12
PCT-US00-33549-27
Sequence 27, Application PC/TUS0033549
GENERAL INFORMATION:
APPLICANT: Sette, Alessandro
APPLICANT: Sidney, John
APPLICANT: Southwood, Scott
APPLICANT: Chesnut, Robert
APPLICANT: Celis, Esteban
APPLICANT: Grey, Howard M.
APPLICANT: EpiImmune Inc.
TITLE OF INVENTION: Inducing Cellular Immune Responses to Human
TITLE OF INVENTION: Papillomavirus Using Peptide and Nucleic Acid
FILE REFERENCE: 018623-016110PC
CURRENT APPLICATION NUMBER: PCT/US00/33549
CURRENT FILING DATE: 2000-12-11
PRIOR APPLICATION NUMBER: US 60/172,705
PRIOR FILING DATE: 1999-12-10
PRIOR APPLICATION NUMBER: US 09/641,528
PRIOR FILING DATE: 2000-08-15
NUMBER OF SEQ ID NOS: 86
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 27
LENGTH: 455
TYPE: PRT
ORGANISM: Human papillomavirus type 11
PCT-US00-33549-27

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: October 17, 2003, 10:47:38 ; Search time 14.2872 Seconds  
(without alignments)  
1517.395 Million cell updates/sec

Title: US-08-913-644-4

Perfect score: 2405  
Sequence: 1 MVSHRAARRKRAVTDLYKT.....FLPKRRRVFFADGVA 461

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	2317.5	96.4	VL2_HPV18	P06793 human papill
2	1973	82.0	VL2_HPV45	P36781 human papill
3	1736	71.4	VL2_HPV45	P27965 human papill
4	1710.5	71.1	VL2_HPV70	P50801 human papill
5	1686.5	70.1	VL2_HPV39	P24839 human papill
6	1550	64.4	VL2_HPV29	P50800 human papill
7	1532.5	63.7	VL2_HPV10	P36747 human papill
8	1530	63.6	VL2_HPV30	P36786 human papill
9	1529	63.6	VL2_HPV53	P36784 human papill
10	1515	63.0	VL2_HPV28	P50799 human papill
11	1513	62.9	VL2_HPV03	P36744 human papill
12	1493.5	62.1	VL2_HPV66	P080960 human papill
13	1486.5	61.8	VL2_HPV56	P36784 human papill
14	1483.5	61.7	VL2_HPV26	P26539 human papill
15	1466.5	61.0	VL2_HPV51	P25487 human papill
16	1457	60.6	VL2_HPV2A	P36784 human papill
17	1423.5	59.2	VL2_HPV27	P36784 human papill
18	1387.5	57.7	VL2_HPV57	P22164 human papill
19	1384	57.5	VL2_HPV61	P080953 human papill
20	1164.5	48.4	VL2_HPV31	P17389 human papill
21	1158	48.1	VL2_HPV13	P00225 human papill
22	1149.5	47.8	VL2_HPV11	P04013 human papill
23	1143.5	47.5	VL2_HPV55	P080939 human papill
24	1139.5	47.4	VL2_HPV44	P080918 human papill
25	1138.5	47.3	VL2_HPV4	P22165 rhesus papill
26	1137	47.3	VL2_HPV6A	P084287 human papill
27	1133.5	47.1	VL2_HPV52	P36783 human papill
28	1133.5	47.1	VL2_HPV16	P03107 human papill
29	1131	47.0	VL2_HPV6B	P03106 human papill
30	1122.5	46.7	VL2_HPV4	P081023 human papill
31	1117	46.4	VL2_HPV42	P27235 human papill
32	1116.5	46.4	VL2_HPV58	P26538 human papill
33	1114	46.3	VL2_HPV33	P06418 human papill

34	1110	46.2	463	1	VL2_PCPV1	O02276 pygmy chimp
35	1108.5	46.1	456	1	VL2_HPV07	P36745 human papill
36	1088	45.2	472	1	VL2_HPV34	P36758 human papill
37	1087.5	45.2	476	1	VL2_HPV32	P36757 human papill
38	1069	44.4	467	1	VL2_HPV40	P36760 human papill
39	1060	44.1	469	1	VL2_HPV35	P27234 human papill
40	667	27.7	519	1	VL2_HPV23	P50797 human papill
41	642.5	26.7	519	1	VL2_HPV14	P36749 human papill
42	640.5	26.6	527	1	VL2_HPV38	P080912 human papill
43	639.5	26.6	533	1	VL2_HPV09	P36746 human papill
44	639	26.6	534	1	VL2_HPV37	P080905 human papill
45	635	26.4	524	1	VL2_HPV22	P50796 human papill

## ALIGNMENTS

RESULT 1  
VL2\_HPV18 STANDARD; PRT; 462 AA.  
AC P06793;  
DT 01-JAN-1988 (Rel. 06, Created)  
DT 01-JAN-1988 (Rel. 06, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Minor capsid protein L2.  
GN L2.  
OS Human Papillomavirus type 18.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10582;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=87283882; PubMed=3039146;  
RA Cole S.T., Danos O.;  
RT "Nucleotide sequence and comparative analysis of the human papillomavirus type 18 genome. Phylogeny of papillomaviruses and repeated structure of the E6 and E7 gene products."  
RL J. Mol. Biol. 193:599-608(1987).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to [license@isb.ch](mailto:license@isb.ch)).  
CC EMBL; X05015; CA28670.1; -.  
DR PIR; B26251; P2WL18.  
DR InterPro; IPR000784; Late\_L2.  
DR Pfam; PF00513; late\_protein\_L2; 1.  
KW Coat protein; late protein.  
SQ SEQUENCE 462 AA; 49596 MW; 6482C186CA3D7E4 CRC64;

Query Match 96.4%; Score 2317.5; DB 1; Length 462;  
Best Local Similarity 96.8%; Pred. No. 3.1e-135;  
Matches 447; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 1 MVSHRAARRKRAVTDLYKTCKSGCPSPVYKVEETTLADITLQMSIGITLGIGIG 60  
DB 1 MVSHRAARRKRAVTDLYKTCKSGCPSPVYKVEETTLADITLQMSIGITLGIGIG 60  
QY 1 TSGGTGRTGYILDGSRNTVVDGPTRPVYIEPVGPTPSTVTLIEDSVYSGAPRP 120  
DB 1 TSGGTGRTGYILDGSRNTVVDGPTRPVYIEPVGPTPSTVTLIEDSVYSGAPRP 120  
QY 121 TFGTSGFDITSGAGTTTAVLDITPSSSTVSISTNFTNPAFSDPSIIEVPGGEVAGNV 180  
DB 121 TFGTSGFDITSGAGTTTAVLDITPSSSTVSISTNFTNPAFSDPSIIEVPGGEVAGNV 180  
QY 181 FVGTPTSGTHGYEIPLOTFASSGTGEEPTSSPLTVRRVAPRLYSRAVQVSYANPE 240  
DB 181 FVGTPTSGTHGYEIPLOTFASSGTGEEPTSSPLTVRRVAPRLYSRAVQVSYANPE 240

[illegible]

```

RESULT 2
VL2_HPVA5
ID      VL2_HPVA5      STANDARD:      PRT:      463 AA.
AC      P36761.
DT      01-JUN-1994 (Rel. 29, Created)
DT      01-JUN-1994 (Rel. 29, Last sequence update)
DT      01-OCT-1996 (Rel. 34, Last annotation update)
DE      Minor capsid protein L2.
GN      L2.
OS      Human papillomavirus type 45.
OC      Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC      Papillomavirus.
OX      NCBI_TaxID=10593;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=94265501; PubMed=8205638;
RA      Delius H., Hofmann B.;
RT      "Primer-directed sequencing of human papillomavirus types.";
RL      Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC      or send an email to license@isb-slb.ch).

```

DR EMBL: X74479; CAA52577.1; -.  
DR PIR: S36565; S36565.  
DR InterPro: IPR000784; late\_L2,  
DR Pfam: PF00513; late\_protein\_L2, 1.  
DR KW Coated protein; Late protein.  
SQ SEQUENCE 463 AA; 49769 MW; 9BD69C97A9D005C6 CRC64;  
Query Match 82.0%; Score 1973; DB 1; Length 463;  
Best Local Similarity 81.7%; Pred. No. 4.1e-114;  
Matches 379; Conservative 33; Mismatches 46; Indels 4; Gaps 3;

```
QY 1 MYSHAAARRKRAVYDLKTKCKROSTGCSDDYVNAKEGTTLAKLIIOWMSLGFIFGLGIG 60
Db 1 MYSHAAARRKRAVADLYRTCKOSGTCPDPYVNAKEGTTLAKLIIOWMSLGFIFGLGIG 60
QY 61 TGSGTGRTGYIPLGGRSNTVVDVGPTRBPVYIEBPDPDSIVTLIEBSSVVTSGAPRP 120
Db 61 TGSGGGRTGYVPLGGRSNTVVDVGPTRBPVYIEBPDPDSIVTLVEDSSVVASGAPRP 120
QY 121 TFGTSGSDITISAGTTTPAVLIDITPSSISVSISTNFTNPASDPSIIEVPOTGEYSQAV 180
Db 121 TFGTSGSEITISSGTTTPAVLIDITPVSISVSISTNFTNPASDPSIIEVPOTGEYSQAV 180
QY 181 FVGPTSGTHYEELIPLQTFASSGGEPEISSTPLPYRVRVGPRLYSAYQOVSAYANE 240
Db 181 FVGPTSSSHYEELIPLQTFASSSGGTEPISSTPLPYRVRVGPRLYSAYQOVSAYQOVS 240
```

Oy	241	FLRPSLLTYDNPRAEPEDOTLLTEPNSWSDSOFMOIIILHAPLATSRRGYAFRLG	300
Db	241	FLRHPSSVLYFDNPRAEPEDOTLLTEPNSWSDSOFMOIIILHAPLATSRRGYAFRLG	3000
Oy	301	ORATMFTRSGTOIGANVHEHDIISPIASPEYIELOPLVSATEDNGLEFDIYADDIDPAMP	360
Db	301	ORATMFTRSGTOIGANVHEHDIISPIASPEYIELOPLVSATEDNGLEFDIYADDIDPAMP	359
Oy	361	VPBRPTTSSAVSYRPTTIS--ASSXSNTVPLTSSMWVRYTGGDTLLP-ITSW#IYS	417
Db	361	VPBRPTTSSAVSYRPTTIS--ASSXSNTVPLTSSMWVRYTGGDTLLP-ITSW#IYS	419
Oy	418	PTAPASTOYIGINGTAYVLMPLYYEIPKKRRKRVPEFADGCVAA	461
Db	418	PTAPASTOYIGINGTAYVLMPLYYEIPKKRRKRVPEFADGCVAA	463

RESULT 3				
ID	VL2_HPVM8	STANDARD;	PRT;	469 AA.
AC	P27965;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Minor capsid protein L2.			
GN	L2.			
OS	Human papillomavirus type ME180.			
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;			
OC	Papillomavirus.			
OX	NCBI_TaxID=10602;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91374616; Pubmed=1716694;			
RA	Reuter S., Delius H., Kahn T., Hofmann B., Zur Hausen H., Schwarz E.;			
RT	"Characterization of a novel human papillomavirus DNA in the cervical			
RT	carcinoma cell line ME180.";			
RL	J. Virol. 65:5564-5568(1991).			

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by, and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M73258; AAF14009.1; -.  
DR PIR; A40509; P2WILPR.  
DR InterPro; IPR000784; Late\_I2.  
DR Pfam; PF00513; late\_protein\_I2; 1.  
KW Coat protein; late protein.  
SQ SEQUENCE 469 AA; 50008 MW; F82425277C9EE9B1 CRC64;

Query Match	71.4%;	Score 1716;	DB 1;	Length 469;
Best Local Similarity	70.1%;	Pred. No. 2.4e 98;		
Matches 329; Conservative	62;	Mismatches 68;	Indels 10;	Gaps 4;

```
QY 1 MSHRARARRRRASVYDLDYTKQSGSTGCSDDVKNVYEGTTLADKLLQWSSLGIFLGGIGIG 60
Db 1 MSHRARARRRRRASAFELXYTKQSGSTGCPDDVINKEVGTTLADKLLQWTSLGIFLGGIGIG 60
QY 61 TSGSGTGRRGYIPLGGRSNTVVDVGPRTPPVIEPVGTGDPISYITLLEDSVYVSGAPRP 120
Db 61 TSGSGTGRRGYIPLGGKPRVYVDVSPARPVYIEPVGTETSIQVLVEDSSVYITSGTVPV 120
QY 121 TETGTSGFDTISAGTTPPAVLDIRPSSSTVSISITTFNTPNAPSDPSIIEVPQGEVSGNV 180
Db 121 TETGTSGFELTSSSTTPPAVLDIRPSSSGVQVSSSTFNAPAFDPITIEVPQGEVSGNV 180
QY 181 FVGTPTSGHNGHEELPLQTFASSGGGEPPISITPLPYRRKAGRRLSRAYQOVSVAANE 240
Db 181 FVSTPTSGHNGHEELPMOVFAHNGTGPDISSTPLPGSRVAGRRLSRADQOVRANSND 240
```

[illegible]

ID	VL2_HP.V70	STANDARD:	PRT:	466 AA.
AC	P50801;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	Minor capsid protein L2.			
GN	L2.			
OS	Human papillomavirus type 70.			
OC	Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;			
OC	Papillomavirus.			
OX	NCHI_TaxID=39457;			
RX	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96249586; PubMed=8815087;			
RA	Forslund O., Hansson B.G.;			
RT	*Human papillomavirus type 70 genome cloned from overlapping PCR			
RL	products: complete nucleotide sequence and genomic organization.,"			
	J. Clin. Microbiol. 34:802-809(1996).			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
DR	EMBL; U21941; AAC54856.1; -;			
DR	InterPro; IPR000784; Late_L2.			
DR	Pfam; PF00513; late_protein_L2; 1.			
DR	Coat protein; Late protein.			
SO	SEQUENCE 466 AA; 49904 MW; 91E15579035F2377 CRC64;			

Query Match	71.1%	Score 1110.5	DB 1	Length 466
Best Local Similarity	71.5%	Pred No.5.2e-98		
Matches	336	Conservative	50	Mismatches 69; Indels 15; Gaps 6;
QY	1	MVSHRAARRKRA	SVTLDKYTCQSGQSCPDVYNNKVEYGTTLADKLQWSSLGFI	LAGLIG 60
DB	1	MVSSRASRRKRAS	ADIKYTCQSGQSCPDVYNNKVEYGTTLADRLQWASLGFI	LAGLIG 60
QY	61	TGSGCGGCTGYI	PLAGRSNTVDVDPMPVPVIEVGYGTDSTVYLIEDSSVYSGARP	120
DB	61	TGTGGGRTGII	PLAGRPSTVDVDPARPVPVIEVGYGTEPSIVQLVRESSVSSGTIP	120
QY	121	TFTGTSGFDIT	SAGTTTTPAVLIDITPSSSTSISITNTFNPAFSDPSLIEVPQTGEVSGN	180
DB	121	TFTGTSGGEIT	SSATTTTPAVLIDITPSSASVQISTSTSYNPAFADPSLIEVPQTGEVSGNI	180
QY	161	FVCGPSTTHGCEET	PLQTFASSGGEPEPISITPLPYRRVAGPLYSRAAQVSYANPE	240
DB	161	FVTTPTSGIHGIEE	LPQVFASSHGTETPISITPQVGSRAAGPLYSRAAHQVAVNPPD	240
QY	241	FLTRPSSLITITDNP	AFEPVDITLLTFEPESNPVDDSPFMDIIRLHRPALTSRKGTVAFSRIG	300

```

Db      241  FVTRPSSEVFEDNPAEPBGDTSLTEFPADTADPDPLDVLRLHRAALTSRRGTFRSRLG 3000
OY      301  QRATMFRSGTOIGARHVFHNDISPAPSEYIELOPVASATNEDGLDIYAD-DIDPAM 3599
Db      301  KKATMFRSGTOIGAOVHYTHDINSNT-ATEDIEMOPLTSTGGLDIYADADIDNAM 3599
OY      360  PVBSRPTSSAVST-----YSPPTISS--ASSYSNVWVPLNSMVDVAVYQPDITLPTS 4111
Db      360  ----LHTTSHTSGTGRSHLSPFSLPSIYSTVIXISNTTTPFTTSMIDIPATTPGPDVLPLPAS 4155
OY      412  V-WPIVSPAPASTOYIGIHGTHVYLMPLVYFPIPKRRKRVPPEFADGVA 460
Db      416  PNLFPVPPPTSIDTTVAIAIOGSNTYLLPLLYFLKKRRKRIPEFTDGVVA 465

```

```

RESULT 5
VL2_HPV39
ID      VL2_HPV39      STANDARD:      PRF:      470 AA.
AC      P24839;
DT      01-MAR-1992 (Rel. 21, Created)
DT      01-MAR-1992 (Rel. 21, Last sequence update)
DT      01-OCT-1996 (Rel. 34, Last annotation update)
DE      Minor capsid protein L2.
GN
OS      Human papillomavirus type 39.
OC      Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC      Papillomavirus.
OX      NCBI_TaxID=10588;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=91135017; PubMed=1847266;
RA      Volpers C., Streeck R.E.;
RT      "Genome organization and nucleotide sequence of human papillomavirus
RT      type 39."
RL      Virology 181:419-423(1991).
-----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC
DR      EMBL: M62849; AAA47055.1; -.
DR      PIR: G38502; P2WT39.
DR      InterPro: IPR000784; Late_L2.
DR      Pfam: PF00513; Late_protein_L2; 1.
KW      Coat protein; Late protein.
SQ      SEQUENCE 470 AA; 50219 MW; 6F07644BE431AD6B CRC64;

```

Query Match	70.1%	Score 1866.5	DB 1	Length 470
Best Local Similarity	69.6%	Pred. No. 1.6e-96		
Matches	328	Conservative 60	Medium 70	Indels 13
				Gaps 7
QY	1	MVSHRAARRKRAVYDLYKTKCKSGSTCDVYKVEGTTIADKIILQWSSIGLFI	60	
Db	1	MVSHRAARRKRAVYDLYKTKCKSGSTCDVYKVEGTTIADKIILQWSSIGLFI	60	
QY	61	TGSGTGGRTGYPLDGRSNTVYDVGPTRPVPIIEVGPEDDSIVTLIEDSSVYTS	120	
Db	61	TGSGTGGRTGYPLDGRSNTVYDVGPTRPVPIIEVGPEDDSIVTLIEDSSVYTS	120	
QY	121	TFTGSGHDITSAGTTTTPAVLIDITPSSTSVSISTENFNPNAPSDPSIIIEV	180	
Db	121	TFTGSGHDITSAGTTTTPAVLIDITPSSTSVSISTENFNPNAPSDPSIIIEV	180	
QY	181	FVSTPTSGTHGEELPMVEFATHGCTEPISTTPPGISRAVGPLYSAHQQVAVS	240	
Db	181	FVSTPTSGTHGEELPMVEFATHGCTEPISTTPPGISRAVGPLYSAHQQVAVS	240	
QY	241	FLTRSSLIITYNPAAFEVDITLTPESNSVYDSDPFMDIIRLRPALTSRRCTV	300	
Db	241	FLTRSSLIITYNPAAFEVDITLTPESNSVYDSDPFMDIIRLRPALTSRRCTV	300	

```

Db      241 EYTHSSFEYFNPAFEPPDTLLTLEADLADPPDLIVLRHRLSRKGTFRSRIG 300
Qy      301 ORATMETRSQTIGARVHFYHDISPAPSEYIELQPLVSA---TEDNGLEIYAD-DID 356
Db      301 KATVYTRGQTIGAOVHYHDISSLAPE-ESIELQPLVHAEPDSADLFEIYADVNN 359
Qy      357 PAMPYPSRRTTSANVTS---PITSSASS--YSNVYPLVLSMDVPYTGPDITLPT 410
Db      360 TYLDIAFNNTKDSG--TYYNTGSLPSVASSATKYANTLIPFSTSNMMPVNTGPDIALPST 418
Qy      411 SWMPYVSPAPASTOY-IGIHGTHYLLMPLYFYFPRKRRVYFPFADGFA 460
Db      419 TPQLPLVPSGPIDTYYAITIGSNYYLLPLFLFKRRIRYFFSDGYA 469

RESULT 6
VL2_HPV29          STANDARD:      PRT: 473 AA.
AC      P50800:
DT      01-OCT-1996 (rel. 34, Created)
DT      01-OCT-1996 (rel. 34, Last sequence update)
DT      01-OCT-1996 (rel. 34, Last annotation update)
DE      Minor capsid protein L2.
GN      L2.
OS      Human papillomavirus type 29.
OC      Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC      Papillomavirus.
OX      NCBI_TaxID=37112;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Delius H.;
RL      Submitted (OCT-1995) to the EMBL/Genbank/DBJ databases.
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; U31784; AAA9434.1; -.
DR      InterPro: IPR000784; Late.L2.
DR      Pfam: PF00513; late.protein.L2; 1.
KW      Coat protein; Late protein.
SQ      SEQUENCE 473 AA; 50631 MW; 70373961C694DB99 CRC64;

Query Match          64.4%; Score 1550; DB 1; Length 473;
Best Local Similarity 63.9%; Pred. No. 3.7e-88;
Matches 304; Conservative 67; Mismatches 85; Indels 20; Gaps 10;

Qy      1 MWSHRAARRKRAASVTDLYKTKQSGTCDVYNNKVEGTTLADKILQWSSLGIFLGLGIG 60
Db      1 MVAHARRRRKRASATOLYRTCKASGTCPDVPYKVEGTTLADRILOWSSLGIVLGLGIG 60
Qy      61 TGSNGRTGYIPLGGRNTVVDVG-PTRRPVYIEPVGPDPSPSYTLIEDSSVYTSQAPR 119
Db      61 TGSNGRTGYIPLGGRNTVVDVG-PTRRPVYIEPVGPDPSPSYTLIEDSSVYTSQAPR 119
Qy      120 PTFGTSGFDITSAGTTTPAVLDTIPSSSTVSISTNTNPAFSDPSIIEYPOGEVSGN 179
Db      121 PTFGTSGFEITTSATTTTPAVLDTIPADNVYISTNNPLTFPSPSILEITPOIGETSGR 180
Qy      180 VFGVTPISGTHGYEIEPIQOTPASSGTGEPISSPPLPVVRVAGFRILSRAYOOVSVANP 239
Db      181 VLVGTPISGTHGYEIEPIQOTPASSGTGEPISSPPLPVVRVAGFRILSRAYOOVSVANP 240
Qy      240 EFLRPSSLLIYDNPAPFVPVTTLFE--PRSNVPDSDPMDIIRLHRLPALTSRRGYRF 296
Db      241 AFLTQPSSEVTFEDNVPDDEDETIIFERPSPGTRVDPDPMIDIVLRHRLPALTSRRGYRF 300
Qy      297 SRLGQATMTFRSGTQIGARVHFYHDISPAPSEYIELQPLV---SATEDNGLEIYAD- 352

```

```

Db      301 SVVQKFSMRKTSRSGTNIGARVHYHDLSPILPT-EDIELEPLLPPADPAEESLVDIYAD 359
Qy      353 -DDIPAMPVPSRPTTS-----SAVSYPTISASSASVYVPLVLSMDVPYTGPDITL 407
Db      360 VDEADMAFGGGRGATTTYGRTTPSVFSTLS--TRIGNVYTLPEWSPVDVPLHTGPDITL 417
Qy      408 PPTSWMPYVSPAPA-STOYIGIHGTHYLLMPLYFYFPRK--RKRPVFFADGFA 460
Db      418 PSSAQMPFV-PVAPADTHYHYIDGADVFLMVEVFPVSRKRRKRLSFLADGFA 472

RESULT 7
VL2_HPV10          STANDARD:      PRT: 470 AA.
AC      P36747;
DT      01-JUN-1994 (rel. 29, Created)
DT      01-JUN-1994 (rel. 29, Last sequence update)
DT      01-OCT-1996 (rel. 34, Last annotation update)
DE      Minor capsid protein L2.
GN      L2.
OS      Human papillomavirus type 10.
OC      Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC      Papillomavirus.
OX      NCBI_TaxID=10603;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE-94265501; PubMed-8205838;
RA      Delius H., Hofmann B.;
RT      "Primer-directed sequencing of human papillomavirus types.";
RL      Curr. Top. Microbiol. Immunol. 186:13-31(1994).
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; X74465; CAA52493.1; -.
DR      PIR: S36536; S36536.
DR      InterPro: IPR000784; Late.L2.
DR      Pfam: PF00513; late.protein.L2; 1.
KW      Coat protein; Late protein.
SQ      SEQUENCE 470 AA; 50583 MW; 056EB847259436E6 CRC64;

Query Match          63.7%; Score 1532.5; DB 1; Length 470;
Best Local Similarity 63.9%; Pred. No. 4.3e-87;
Matches 301; Conservative 69; Mismatches 88; Indels 13; Gaps 9;

Qy      1 MWSHRAARRKRAASVTDLYKTKQSGTCDVYNNKVEGTTLADKILQWSSLGIFLGLGIG 60
Db      1 MVAQARRRRKRASATOLYRTCKASGTCPDVPYKVEGTTLADRILOWSSLGIVLGLGIG 60
Qy      61 TGSNGRTGYIPLGGRNTVVDVG-PTRRPVYIEPVGPDPSPSYTLIEDSSVYTSQAPR 119
Db      61 TGSNGRTGYIPLGGRNTVVDVG-PTRRPVYIEPVGPDPSPSYTLIEDSSVYTSQAPR 119
Qy      120 PTFGTSGFDITSAGTTTPAVLDTIPSSSTVSISTNTNPAFSDPSIIEYPOGEVSGN 179
Db      121 PTFSTSEFVYTSATTTTPAVLDTIPASENVYISTNTNPAFTFSPSILEYVPOGEVSGH 180
Qy      180 VFGVTPISGTHGYEIEPIQOTPASSGTGEPISSPPLPVVRVAGFRILSRAYOOVSVANP 239
Db      181 ILISPTAGTHGYEIEPIQOTPASSGTGEPISSPPLPVVRVAGFRILSRAYOOVSVANP 240
Qy      240 EFLRPSSLLIYDNPAPFVPVTTLFE--PRSNVPDSDPMDIIRLHRLPALTSRRGYRF 296
Db      241 AFLRPSSLLIYDNPAPFVPVTTLFE--PRSNVPDSDPMDIIRLHRLPALTSRRGYRF 300
Qy      297 SRLGQATMTFRSGTQIGARVHFYHDISPAPSEYIELQPLV---SATEDNGLEIYAD- 354

```

DB 301 SRLGQKFSMRTSGKIGARVHYQDLSPAPI-EDIEMEPILAPASDITVIDFADVD 359

QY 355 IIPAMPVPSRPTSS-AVSTYSPTISSASS-YSNVTPLTSSMDVPTGPDITLPTSY 412

DB 360 GVAFAFEGYRSTQSGYNTTSLSTLSTKYGNVITPFVS PVDVTLHGTGPDIVLPTSAQ 419

QY 413 WPIVSTAPAR-SVOYIGIHGTHYIMPLPYFIP-KKKRVPYFEPADGFAA 460

DB 420 WPIV-PLSPADTTHYVYIDGDEYIMLPVTFHFSRRRRRRKRVSYFADGTLA 469

RESULT 8

VL2\_HPV30 STANDARD; PRT; 463 AA.

AC P36756;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Minor capsid protein L2.

GN L2.

OS Human papillomavirus type 30.

OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

OC Papillomavirus.

OX NCBI\_Taxid=10611;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94265501; PubMed=8205838;

RA Delius H., Hofmann B.;

RT "Primer-directed sequencing of human papillomavirus types.";

RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X74474; CAA52547.1; -

DR PIR; S36507; S36507.

DR InterPro: IPR000784; Late.L2.

DR Pfam: PF00513; late\_protein\_L2; 1.

KW Coat protein; late protein.

SQ SEQUENCE 463 AA; 49903 MW; 9B5B526107F2B072 CRC64;

Query Match 63.6%; Score 1530; DB 1; Length 463;

Best local Similarity 63.2%; Pred. No. 6, 1e-87;

Matches 295; Conservative 68; Mismatches 94; Indels 10; Gaps 8;

QY 1 MVSHRAARRRRASVTDLYTKCKSGTCPSDVNKNVKGTTLADKILQWSSLGIFLAGIGIG 60

DB 1 MVAHRARRRRRRASATQLYCTCKQAGTCPSDVINKIEHTTLADKILQWSSLGIFLAGIGIG 60

QY 61 TSGGTGRCGYITPLGGRSNVVDVGPTRPVVIEPVGPDPSIVTLIEDSSVYTSQAPRP 120

DB 61 TGGSGGRCGYITPLGGRSNVVDVGPTRPVVIEPVGPDPSIVTLIEDSSVYTSQAPRP 120

QY 121 TFGTSGFDITSGAGTTTAVLDITPSTSVISSTNTFNPAFSDPSIIIEVPGTGEVSGNV 180

DB 121 NFGTSGFVETSSSTTTTAVLDITPSTSVISSTNTFNPAFSDPSIIIEVPGTGEVSGNV 180

QY 181 FVGTPTSGHGEIPLQTFPAGSGTEPISSTPLTVRRVAPRLYSRAVOVSYANPE 240

DB 181 LVSTPTSGVHGEIPLQTFPAGSGTEPISSTPLTVRRVAPRLYSRAVOVSYANPE 240

QY 241 FLTRPSLLIYDNDAPAPVDTTLTFEPRSNVPSDPMDDIIRLRPALTSRGRVRSRLG 300

DB 241 FLTRPSLLIYDNDAPAPVDTTLTFEPRSNVPSDPMDDIIRLRPALTSRGRVRSRLG 300

QY 301 ORATMFTSGTQIGARVHYNDISPLAPSEYIELOPLVSATED--GLFDIYADDIDP 359

DB 301 TKATMFTSGKQIGARVHYNDISPLAPSEYIELOPLVSATED--GLFDIYADDIDP 359

QY 360 PVPSRPT--TSSAVSYSTPSSASSYSNVTPLTSSMDVPTGPDITLPT--PTGVMPDI 415

DB 359 PVSSHLSIATPSRLPTNTVPLSFSQNTNTVITPLGKWDVPIVSGDVIPLPTPTT-WP- 416

QY 416 VSPFAPASROY-IGIHGTHYIMPLPYFIPKKRRRPYFPAADGFAA 461

DB 417 YAPQADPTTHVDVYIHGTFALMPVYFLRRRRRRKHVYFLADGVAA 463

RESULT 9

VL2\_HPV53 STANDARD; PRT; 463 AA.

AC P36764;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Minor capsid protein L2.

GN L2.

OS Human papillomavirus type 53.

OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

OC Papillomavirus.

OX NCBI\_Taxid=10619;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=94265501; PubMed=8205838;

RA Delius H., Hofmann B.;

RT "Primer-directed sequencing of human papillomavirus types.";

RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; X74482; CAA52594.1; -

DR PIR; S36530; S36530.

DR InterPro: IPR000784; Late.L2.

DR Pfam: PF00513; late\_protein\_L2; 1.

KW Coat protein; late protein.

SQ SEQUENCE 463 AA; 50118 MW; F065B5E7A518EA3E CRC64;

Query Match 63.6%; Score 1529; DB 1; Length 463;

Best local Similarity 63.2%; Pred. No. 7e-87;

Matches 297; Conservative 63; Mismatches 94; Indels 16; Gaps 9;

QY 1 MVSHRAARRRRASVTDLYTKCKSGTCPSDVNKNVKGTTLADKILQWSSLGIFLAGIGIG 60

DB 1 MVAHRARRRRRRASATQLYCTCKQAGTCPSDVINKIEHTTLADKILQWSSLGIFLAGIGIG 60

QY 61 TSGGTGRCGYITPLGGRSNVVDVGPTRPVVIEPVGPDPSIVTLIEDSSVYTSQAPRP 120

DB 61 TGGSGGRCGYITPLGGRSNVVDVGPTRPVVIEPVGPDPSIVTLIEDSSVYTSQAPRP 120

QY 121 TFGTSGFDITSGAGTTTAVLDITPSTSVISSTNTFNPAFSDPSIIIEVPGTGEVSGNV 180

DB 121 NFGTSGFVETSSSTTTTAVLDITPSTSVISSTNTFNPAFSDPSIIIEVPGTGEVSGNV 180

QY 181 FVGTPTSGHGEIPLQTFPAGSGTEPISSTPLTVRRVAPRLYSRAVOVSYANPE 240

DB 181 LVSTPTSGVHGEIPLQTFPAGSGTEPISSTPLTVRRVAPRLYSRAVOVSYANPE 240

QY 241 FLTRPSLLIYDNDAPAPVDTTLTFEPRSNVPSDPMDDIIRLRPALTSRGRVRSRLG 300

DB 241 FLTRPSLLIYDNDAPAPVDTTLTFEPRSNVPSDPMDDIIRLRPALTSRGRVRSRLG 300

QY 301 ORATMFTSGTQIGARVHYNDISPLAPSEYIELOPLVSATED--GLFDIYADDIDP 357

DB 301 TKATMFTSGKQIGARVHYNDISPLAPSEYIELOPLVSATED--GLFDIYADDIDP 357

QY 358 AMPVPSR--PTTSSAVSTYPTTSSASYSVNVPLSSMDVPVYTGDTL---PTSV 412  
DB 357 EAPVSSRSISATNRLPTNTPVLSFSGSTSNVTIPFGSDMDVPYISGDDVPLPGPP-- 414  
OY 413 WPVSPRPASTOY-IGIGHGTHYMLPLYFLPKRRKRVPEFADGFVA 461  
DB 415 WPFOSPFDTTHDVYIGSTFALMPVFLKRRRRKRIPTFLADGVAA 463

RESULT 10  
ID V12\_HPV28 STANDARD: PRT: 473 AA.  
AC P50799;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Minor capsid protein L2.  
GN L2.  
OS Human papillomavirus type 28.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=37111;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Delius H.;  
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: U31783; AAA79427.1; -  
DR InterPro: IPR000784; Late\_L2.  
DR Pfam: PF00513; late-protein\_L2; 1.  
DR Coated protein; Late protein.  
KW Coated protein; Late protein.  
SQ SEQUENCE 473 AA; 50490 MW; E6461BF2D333089C CRC64;

Query Match 63.0%; Score 1515; DB 1; Length 473;  
Best local similarity 62.4%; Pred. No. 5.2e-86;  
Matches 296; Conservative 71; Mismatches 91; Indels 16; Gaps 10;

QY 1 MVSRAARRKRASVYDLKYTKCKOSGCPSDVNVKVEGTTLADKILQWSSLGIFLAGLGIG 60  
DB 1 MVARARRRRKRASATQLYRCKAAGTCPPDVIPRYEGTTLADRLQWGGIGLYLGIGIG 60  
OY 61 TSGGTGRTGYIPLAGKSNIVYDVG-PTRPPVYIEPVGPTDPSIVTLIEDSSVVTSGAPR 119  
DB 61 TSGGTGRTGYIPLAGKSNIVYDVG-PTRPPVYIEPVGPTDPSIVTLIEDSSVVTSGAPR 119  
OY 120 PTFGTSGFDITSGTTTTPALDITPSTSVSISTNTNPAFSDPSIIEVPOGVEGSGN 179  
DB 121 PTFGTSGFEVTSATTTTPALDITPASNIVSSNTNPAFTEPISLLEVPQNGEVSGH 180  
OY 180 VFVGTPTSGTHGYEIEIPLQTFPAGSGTEPISTPLPTVRRVAGPRLYSRAVQOVSANP 239  
DB 181 ILVSTPTAGTHSYEIEIEMETPASPCTGNEPISSITPVGVSKIADPRLYAKAVTYQKATDP 240  
OY 240 EFLTRPSSLITYNDPAFEPVDITLFE--PRSNVPSDSFMDIIRLRPALTSRRGTVRF 296  
DB 241 AFLTRPSLTIVFNDPAFEPDEETIIFERPYPSPQVDPDFDIIRLRPALTSRRGTVRF 300  
OY 297 SRLGQRATMPTFRSGTOGARVHFHDISPAPSEYIELOPLVATD---NGLEFDIAD 353  
DB 301 SRLGKLSMHTRSKSGIGARVHYQDLSPIGPT-EDIEMEPLAPAEAAVADSISYDVAD 359  
OY 354 --DIDPAMPVPSR--PTTSSAVSTYPTTSSASYSVNVPLSSMDVPVYTGDTLPP 409  
DB 360 VEDADIAFTGSRKATSRGTYVSSPLSSTLTITKGNVTIFPVSPVYVHILHGPDIITPA 419

QY 410 TSWPVISSPAPA-STOYIGIGHGTHYMLPLYFIP--KKRRVPEFADGFVA 460  
DB 420 STOMPFEV-PLVPADTHYVYIDGDFYLMVPLFVPRRRRRKRRLSYFLADGVTA 472

RESULT 11  
ID V12\_HPV03 STANDARD: PRT: 473 AA.  
AC P36744;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Minor capsid protein L2.  
GN L2.  
OS Human papillomavirus type 3.  
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;  
OC Papillomavirus.  
OX NCBI\_TaxID=10614;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94265501; PubMed=8205838;  
RA Delius H.; Hofmann B.;  
RT "Primer-directed sequencing of human papillomavirus types.";  
RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC  
CC EMBL: X74462; CA52473.1; -  
DR PIR: S36553; S36553.  
DR InterPro: IPR000784; Late\_L2.  
DR Pfam: PF00513; late-protein\_L2; 1.  
DR Coated protein; Late protein.  
KW Coated protein; Late protein.  
SQ SEQUENCE 473 AA; 50662 MW; DD8298FE517A04DD CRC64;

Query Match 62.9%; Score 1513; DB 1; Length 473;  
Best local similarity 62.2%; Pred. No. 6.9e-86;  
Matches 295; Conservative 71; Mismatches 92; Indels 16; Gaps 10;

QY 1 MVSRAARRKRASVYDLKYTKCKOSGCPSDVNVKVEGTTLADKILQWSSLGIFLAGLGIG 60  
DB 1 MVARARRRRKRASATQLYRCKAAGTCPPDVIPRYEGTTLADRLQWGGIGLYLGIGIG 60  
OY 61 TSGGTGRTGYIPLAGKSNIVYDVG-PTRPPVYIEPVGPTDPSIVTLIEDSSVVTSGAPR 119  
DB 61 TSGGTGRTGYIPLAGKSNIVYDVG-PTRPPVYIEPVGPTDPSIVTLIEDSSVVTSGAPR 119  
OY 120 PTFGTSGFDITSGTTTTPALDITPSTSVSISTNTNPAFSDPSIIEVPOGVEGSGN 179  
DB 121 PTFGTSGFEVTSATTTTPALDITPASNIVSSNTNPAFTEPISLLEVPQNGEVSGH 180  
OY 180 VFVGTPTSGTHGYEIEIPLQTFPAGSGTEPISTPLPTVRRVAGPRLYSRAVQOVSANP 239  
DB 181 ILVSTPTAGTHSYEIEIEMETPASPCTGNEPISSITPVGVSKIADPRLYAKAVTYQKATDP 240  
OY 240 EFLTRPSSLITYNDPAFEPVDITLFE--PRSNVPSDSFMDIIRLRPALTSRRGTVRF 296  
DB 241 AFLTRPSLTIVFNDPAFEPDEETIIFERPYPSPQVDPDFDIIRLRPALTSRRGTVRF 300  
OY 297 SRLGQRATMPTFRSGTOGARVHFHDISPAPSEYIELOPLV---SATEDNGLEFDIYA- 352  
DB 301 SRVQKLSMHTRSKSGIGARVHYQDLSPIGPT-EDIEMEPLAPAEAAVADSISYDVAD 359  
OY 353 --DIDPAMPVPSR--PTTSSAVSTYPTTSSASYSVNVPLSSMDVPVYTGDTLPP 409  
DB 360 VDDADIGFTSGGRDITLSRGATVSPSSSTLTITKGNVTIFPVSPVYVHILHGPDIITPA 419  
OY 410 TSWPVISSPAPA-STOYIGIGHGTHYMLPLYFIP--KKRRVPEFADGFVA 460



DB 420 SNAWPEV-PLSEVDYTHHYIDGDIYLPVTFELPRRRRRKRVSTFLADGVA 472

RESULT 12

ID V12\_HPV56 STANDARD; PRT; 464 AA.

AC 080360;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE Minor capsid protein L2.

GN L2.

OS Human papillomavirus type 66.

OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

OC Papillomavirus.

OX NCBI\_Taxid=37119;

RN [1]

RP SEQUENCE FROM N.A.

RA Delius H.;

RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: U31794; AA79504.1; -

DR Interpro: IPR000784; Late\_L2.

DR Pfam: PF00513; Late\_protein\_L2; 1.

KW Coat protein; Late protein.

SQ SEQUENCE 464 AA; 49988 MW; 19C3D0FCE131PE2 CRC64;

Query Match

Best Local Similarity 62.1%; Score 1493.5; DB 1; Length 464;

Matches 291; Conservative 63; Mismatches 104; Indels 9; Gaps 8;

QY 1 MYSHRAARRKRAVTDLYKTCOSGCPSPDVNVKVGTTLADKILQWSSLGIFLGGIG 60  
DB 1 MAHRAIRRRKRAATDLYKTCOSGCPSPDVNVKVGTTLADKILQWSSLGIFLGGIG 60  
QY 61 TSGSGGRTGYIPLGSRNTVDVGPTRPVVIEVPPTDPSIVTLIEDSVVTSGAPR 120  
DB 61 TSGSGGRTGYIPLGSRNTVDVGPTRPVVIEVPPTDPSIVTLIEDSVVTSGAPR 120  
QY 121 TFGTSGFDITSGTTTAVLDITPSTSVSISTFTFTNPAPDSIEVPQGEVSGNV 180  
DB 121 TFGTSGFDITSGTTTAVLDITPSTSVSISTFTFTNPAPDSIEVPQGEVSGNV 180  
QY 181 FVGTPTSGHGEYELPQTFASSGTEEPISSTPLTVRRVAGPRISRAYQOVSAPE 240  
DB 181 FVGTPTSGHGEYELPQTFASSGTEEPISSTPLTVRRVAGPRISRAYQOVSAPE 240  
QY 241 FLTRPSSLLTYDNPAPFVDVTLTFEPRNSNVPDSDPMDIIRLRPALTSRGTVRSR 300  
DB 241 FLTRPSSLLTYDNPAPFVDVTLTFEPRNSNVPDSDPMDIIRLRPALTSRGTVRSR 300  
QY 241 FLTRPSSLLTYDNPAPFVDVTLTFEPRNSNVPDSDPMDIIRLRPALTSRGTVRSR 300  
DB 241 FLTRPSSLLTYDNPAPFVDVTLTFEPRNSNVPDSDPMDIIRLRPALTSRGTVRSR 300  
QY 301 ORATMTFRTSGTOIGAVHVFHDIPTAPSEYELDPLVSATDE-NGLEDIYADDDPAM 359  
DB 301 ORATMTFRTSGTOIGAVHVFHDIPTAPSEYELDPLVSATDE-NGLEDIYADDDPAM 359  
QY 360 PVPSPRT--TSSAVSTYSV-TISSASSYGNVTVPLTSSMDVPVYTGPDITLPP-TSVMP 415  
DB 360 PVPSPRT--TSSAVSTYSV-TISSASSYGNVTVPLTSSMDVPVYTGPDITLPP-TSVMP 415  
QY 359 PLSFRSGATPSAQLPIKPSSTLSFASNTANTVAPLGNWETPFYSGPDIVLPFGPSTWPR 418  
DB 359 PLSFRSGATPSAQLPIKPSSTLSFASNTANTVAPLGNWETPFYSGPDIVLPFGPSTWPR 418  
QY 416 VSPFADAS-TQYIGHGHYIYIMPLIYPIPKRRKRVYFFADGVAA 461  
DB 416 VSPFADAS-TQYIGHGHYIYIMPLIYPIPKRRKRVYFFADGVAA 461  
QY 419 V-POSPDVTHDVYIOGATFALMPYVFFRRRRRRRIPIFFADGVAA 464  
DB 419 V-POSPDVTHDVYIOGATFALMPYVFFRRRRRRRIPIFFADGVAA 464

RESULT 13

ID V12\_HPV56 STANDARD; PRT; 464 AA.

AC P36765;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 01-OCT-1996 (Rel. 34, Last annotation update)

DE Minor capsid protein L2.

GN L2.

OS Human papillomavirus type 56.

OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

OC Papillomavirus

OX NCBI\_Taxid=10596;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-9426501; Pubmed-8205838;

RA Delius H., Hofmann B.;

RT "Primer-directed sequencing of human papillomavirus types.";

RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: X74483; CA52599.1; -

DR PIR: S36582; S36582.

DR Interpro: IPR000784; Late\_L2.

DR Pfam: PF00513; Late\_protein\_L2; 1.

KW Coat protein; Late protein.

SQ SEQUENCE 464 AA; 49983 MW; 41B6687D5C8E7E21 CRC64;

Query Match

Best Local Similarity 61.8%; Score 1486.5; DB 1; Length 464;

Matches 284; Conservative 68; Mismatches 108; Indels 5; Gaps 4;

QY 1 MYSHRAARRKRAVTDLYKTCOSGCPSPDVNVKVGTTLADKILQWSSLGIFLGGIG 60  
DB 1 MAHRAIRRRKRAATDLYKTCOSGCPSPDVNVKVGTTLADKILQWSSLGIFLGGIG 60  
QY 61 TSGSGGRTGYIPLGSRNTVDVGPTRPVVIEVPPTDPSIVTLIEDSVVTSGAPR 120  
DB 61 TSGSGGRTGYIPLGSRNTVDVGPTRPVVIEVPPTDPSIVTLIEDSVVTSGAPR 120  
QY 121 TFGTSGFDITSGTTTAVLDITPSTSVSISTFTFTNPAPDSIEVPQGEVSGNV 180  
DB 121 TFGTSGFDITSGTTTAVLDITPSTSVSISTFTFTNPAPDSIEVPQGEVSGNV 180  
QY 181 FVGTPTSGHGEYELPQTFASSGTEEPISSTPLTVRRVAGPRISRAYQOVSAPE 240  
DB 181 FVGTPTSGHGEYELPQTFASSGTEEPISSTPLTVRRVAGPRISRAYQOVSAPE 240  
QY 241 FLTRPSSLLTYDNPAPFVDVTLTFEPRNSNVPDSDPMDIIRLRPALTSRGTVRSR 300  
DB 241 FLTRPSSLLTYDNPAPFVDVTLTFEPRNSNVPDSDPMDIIRLRPALTSRGTVRSR 300  
QY 241 FLTRPSSLLTYDNPAPFVDVTLTFEPRNSNVPDSDPMDIIRLRPALTSRGTVRSR 300  
DB 241 FLTRPSSLLTYDNPAPFVDVTLTFEPRNSNVPDSDPMDIIRLRPALTSRGTVRSR 300  
QY 301 ORATMTFRTSGTOIGAVHVFHDIPTAPSEYELDPLVSATDE-NGLEDIYADDDPAM 359  
DB 301 ORATMTFRTSGTOIGAVHVFHDIPTAPSEYELDPLVSATDE-NGLEDIYADDDPAM 359  
QY 358 AMPVSPRTTSSAVSTYSV-TISSASSYGNVTVPLTSSMDVPVYTGPDITLPP-TSVMP 416  
DB 358 AMPVSPRTTSSAVSTYSV-TISSASSYGNVTVPLTSSMDVPVYTGPDITLPP-TSVMP 416  
QY 360 GLSSQVAVNPASALPIKPSSTLSFASNTANTVAPLGNWETPFYSGPDIVLPFGPSTWPR 419  
DB 360 GLSSQVAVNPASALPIKPSSTLSFASNTANTVAPLGNWETPFYSGPDIVLPFGPSTWPR 419  
QY 417 SPTAPASTOYIGHGHYIYIMPLIYPIPKRRKRVYFFADGVAA 461  
DB 417 SPTAPASTOYIGHGHYIYIMPLIYPIPKRRKRVYFFADGVAA 461  
QY 420 POSPDVTHDVYIOGSSFALMPYVFFRRRRRRRIPIFFADGVAA 464  
DB 420 POSPDVTHDVYIOGSSFALMPYVFFRRRRRRRIPIFFADGVAA 464

RESULT 14

VL2\_HPV26 STANDARD: PRT: 472 AA.

AC P36754; 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DR 01-OCT-1996 (Rel. 34, Last annotation update)

DE Minor capsid protein L2.

GN L2.

OS Human papillomavirus type 26.

OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

OC Papillomavirus.

OX NCBI\_TaxID=31549;

RA MEDLINE=94265501; Pubmed=8205838;

RP SEQUENCE FROM N.A.

RX Delius H., Hotmann B.;

RT Primer-directed sequencing of human papillomavirus types.;

RL Curr. Top. Microbiol. Immunol. 186:13-31(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL; X74472; CA52534.1; -

DR PIR; S36548; S36548.

DR InterPro; IPR000784; Late\_L2.

DR Pfam; PF00513; late\_protein\_L2; 1.

KW Coat protein; Late protein.

SEQUENCE 472 AA; 50613 MW; DA05C41B80FB6C50 CRC64;

Query Match 61.7%; Score 1483.5; DB 1; Length 472;

Best Local Similarity 60.8%; Pred. No. 4.4e-84;

Matches 287; Conservative 81; Mismatches 91; Indels 13; Gaps 10;

QY 1 MVSRRARRRRASVTDLYKTKQSGTCDSDVNVNVEGTTLADKILQMSGLIFGLGIG 60

DB 1 MVAIRARRRRARRASVTDLYKTKKAAGTCPPDVNKRVEGTTLADKILQMSGLIFGLGIG 60

QY 61 TGSCTGRTGYIPLGGSN-TVDVGPTRPVVIEPVPDPSPVTLIEDSSVYTSQAPR 119

DB 61 TGTSGGRTGYIPLGGGRPGVVDIGTRPPIIIEPVPDPSPVTLIEDSSVYTSQAPR 120

QY 120 PTFGTGSGFDITSGTTPPAVLDTTPSSTSVSISTNFTNPAFSDPSIIEVPGTGVSGN 179

DB 121 PTFSGGNGFELTSSATTPAVLDITPSAGTVHVTSTNQNPLYLEPP-IDIPQAGEASGH 179

QY 180 VFTGPTSGTHGYEIPLOTFPSS-SGTGEPISTPLPYRRAVAGPRLYSRAYOVSAN 238

DB 180 IFTTSTAGHSHYEIMEVFASNGTGLEPISSTPIPGIOKVAAPRLYSRAYOVSAN 239

QY 239 PEFLTRPSLLTYNPAFEVDTLTFEPKSNV-PDSDFMDIILHAPALTSRGTGRFS 297

DB 240 PNFIGNSTFTVFNPAPEPIDEITLVASSSTVAPDDFDIILHAPALTSRGTGRFS 299

QY 298 RLQGRATMTFRSGQIGARVHFHYDISPIAPSPY-TELQPLVSATEDNG-LFDIYAD- 353

DB 300 RLQGRATMTFRSGQIGARVHFHYDISPIAPSPY-TELQPLVSATEDNG-LFDIYAD- 359

QY 354 DIDPAMVP--SRPTSSAVSTYSPTISSA--SSYSNVTVPDLSMDVPVYTGPDITLPP 409

DB 360 DYPVSIHTPRKSYSPITLPPKRIASNVFSSHTSTVAVPLSLSSFELPVSSTGSDITPT 419

QY 410 TS-VMPVSPAPASTQYIGIHGHYVLMPLLYYFIPKRRKRVPPFADGEVA 460

DB 420 SSPTWPSLPPPTTNLPAIYVHGDNLYLMPYIYLHKKRRKRPFFSDGEVA 471

VL2\_HPV51 STANDARD: PRT: 468 AA.

AC P26539; 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DR 01-OCT-1996 (Rel. 34, Last annotation update)

DE Minor capsid protein L2.

GN L2.

OS Human papillomavirus type 51.

OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;

OC Papillomavirus.

OX NCBI\_TaxID=10595;

RA MEDLINE=91303675; Pubmed=1649326;

RP SEQUENCE FROM N.A.

RX Lungu O., Crum C.P., Silverstein S.J.;

RT "Biologic properties and nucleotide sequence analysis of human papillomavirus type 51.;"

RL J. Virol. 65:4216-4225(1991).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL; M62877; -; NOT\_ANNOTATED\_CDS.

DR PIR; H40415; P2WL51.

DR InterPro; IPR000784; Late\_L2.

DR Pfam; PF00513; late\_protein\_L2; 1.

KW Coat protein; Late protein.

SEQUENCE 468 AA; 50720 MW; B5BAC02A12712743 CRC64;

Query Match 61.0%; Score 1466.5; DB 1; Length 468;

Best Local Similarity 61.1%; Pred. No. 4.8e-83;

Matches 288; Conservative 78; Mismatches 90; Indels 15; Gaps 10;

QY 1 MVSRRARRRRASVTDLYKTKQSGTCDSDVNVNVEGTTLADKILQMSGLIFGLGIG 60

DB 1 MVAIRARRRRARRASVTDLYKTKKAAGTCPPDVNKRVEGTTLADKILQMSGLIFGLGIG 60

QY 61 TGSCTGRTGYIPLGGSN-TVDVGPTRPVVIEPVPDPSPVTLIEDSSVYTSQAPR 119

DB 61 TGTSGGRTGYIPLGGGRPGVVDIGTRPPIIIEPVPDPSPVTLIEDSSVYTSQAPR 120

QY 120 PTFGTGSGFDITSGTTPPAVLDTTPSSTSVSISTNFTNPAFSDPSIIEVPGTGVSGN 179

DB 121 PTFGTGSGFDITSGTTPPAVLDTTPSAGTVHVTSTNQNPLYLEPP-IDIPQAGEASGH 179

QY 180 VFTGPTSGTHGYEIPLOTFPSS-SGTGEPISTPLPYRRAVAGPRLYSRAYOVSAN 238

DB 180 IYLVHSGTHGYEIMEVFASNGTGLEPISSTPIPGIOKVAAPRLYSRAYOVSAN 239

QY 239 PEFLTRPSLLTYNPAFEVDTLTFEPKSNV-PDSDFMDIILHAPALTSRGTGRFS 297

DB 240 PNFIGNSTFTVFNPAPEPIDEITLVASSSTVAPDDFDIILHAPALTSRGTGRFS 299

QY 298 RLQGRATMTFRSGQIGARVHFHYDISPIAPSPY-TELQPLVSATEDNG-LFDIYADIDP 357

DB 300 RLQGRATMTFRSGQIGARVHFHYDISPIAPSPY-TELQPLVSATEDNG-LFDIYADIDP 356

QY 358 A-----MPVSRPTSSAVSTYSPTISSA--SSYSNVTVPDLSMDVPVYTGPDITLPP 410

DB 357 AETGFIQPTHTTTPMSHSLRQLPSLSSMSSTVAVNTIPFTSPSTVPIHTGPDVVLPTS 416

QY 411 -SWPVPSPAPASTQYIGIHGHYVLMPLLYYFIPKRRKRVPPFADGEVA 460

DB 417 PTWVPVPHNSIDTKHSIVILGIDYILMPYTHLHKKRRKRPFFSDGEVA 467



THIS PAGE BLANK (USPTO)



```

Db      61 TSGGTGRTGYIPLGGRNTIVDVSPAKRPVIEVPGPDSIVTLVEDSSVITSGAPAP 120
Qy      121 TFGTSGFDITAGATTDAVIDITPSSVSISITNTFNPAFSDSIIIEVPGTGEVSGNV 180
Db      121 TFGTSGFEISSTSTTDAVIDITPTS-SVQISSSEFINPAFTDPSVIEVPGTGEISGNI 179
Qy      181 FVGTSTGTHGEELIPLDTFASSGTCEPISSTPLPYARRVAGPRLYSRAVQVSVANPE 240
Db      180 LISTSTGSHGEEELIPMOTFNEGTEGLEPISSTPNPYARRVAGPRLYSRAVQVSVANPE 239
Qy      241 FLTRSSLTITDNPFAFEPVDITLTFEPNSNVPDSDFMIIIRLHRALTSRGTGVRFSRLG 300
Db      240 FLTRSTVTYVYNPAVDIDITLTFDPSSEVDPDPFMDIIRLHRALTSRGTGVRFSRLG 299
Qy      301 ORATMFTSRGTQIGARVHFHYHDISPASPEXIELQPLV---ATEDNGLEDIYADIDP 357
Db      300 ORATMFTSRGTQIGARVHFHYHDISP- PHAEDIELEQPLVSSQAAND--IYDIDIRDE 356
Qy      358 AMPVPSRPTT---SSAVSYSPITSSASVSNTVPLTSSMDVAVYTGPDITLPTPS-V 412
Db      357 A-PISTANTATFTPKSSFOQSLTRSSASTSNVPLATADVNTGPDITLPTNTIV 415
Qy      413 MPVSPAPASTQYIGIHGTHYLLMPLYFIKRRKRVYFPADGEVA 460
Db      416 EPTVSTPTFTTQISINIGTINFLMPLYFPRKRRVYFPPTDGSMA 463

```

## RESULT 2

```

Q9WHG6 PRELIMINARY; PRT; 467 AA.
ID AC Q9WHG6
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE Putative minor capsid protein L2.
GN L2.
OS Human papillomavirus candid HPV85.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=151757;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=20047972; PubMed=10580054;
RA Chow V.T.K., Leong W.F.;
RT "Complete nucleotide sequence, genomic organization and phylogenetic
RT analysis of a novel genital human papillomavirus type, HLT7474-S."
RL J. Gen. Virol. 80:2923-2929(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Chow V.T.K., Leong W.F.;
RL Submitted (Feb-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF131950; AAD24187.1; -.
DR InterPro; IPR000784; Late_L2.
DR Pfam; PF00513; Late_protein_L2; 1.
SQ SEQUENCE 467 AA; 50394 MW; 773F18A2525F5235 CRC64;

```

Query Match 73.8%; Score 1774; DB 12; Length 467;  
 Best Local Similarity 72.4%; Pred. No. 5.9e-111;  
 Matches 343; Conservative 51; Mismatches 58; Indels 22; Gaps 7;

```

Qy      1 MWSHRAARRKRASTYDLYKTKQSGTSPDVYVKNVEGTTTLADKILQWSSLGIFLGIG 60
Db      1 MWSHRAARRKRASTADLYKTKQSGTSPDVYVKNVEGTTTLADKILQWSSLGIFLGIG 60
Qy      61 TSGGTGRTGYIPLGGRNTIVDVSPAKRPVIEVPGTDPISIVTLIEDSSVITSGAPAP 120
Db      61 TSGGTGRTGYIPLGGRNTIVDVSPAKRPVIEVPGTDPISIVTLIEDSSVITSGAPAP 120
Qy      121 TFGTSGFDITAGATTDAVIDITPSSVSISITNTFNPAFSDSIIIEVPGTGEVSGNV 180
Db      121 TFGTSGFEISSTSTTDAVIDITPTS-SVQISSSEFINPAFTDPSVIEVPGTGEISGNI 179
Qy      181 FVGTSTGTHGEELIPLDTFASSGTCEPISSTPLPYARRVAGPRLYSRAVQVSVANPE 240

```

```

Db      181 FITTPTSGTHGEEELIPMHTFAQTGCTEPISSTPLDGRVRAVAGPRLYSQAQVAKITNSD 240
Qy      241 FLTRSSLTITDNPFAFEPVDITLTFEPNSNVPDSDFMIIIRLHRALTSRGTGVRFSRLG 300
Db      241 FLTRSSLTITDNPFAFEPVDITLTFEPNSNVPDSDFMIIIRLHRALTSRGTGVRFSRLG 300
Qy      301 ORATMFTSRGTQIGARVHFHYHDISPASPEXIELQPLV---SATED-NGLEDIYAD-DI 355
Db      301 KLFMTSRGTQIGARVHFHYHDISPISHIGESIEMQPLLPDAVA7ADTNGLEDIYADTDI 360
Qy      356 DPAMVPSR-----PTSSASVSITPTSSAS-YSNVVPLTSSMDVAVYTGPDITL 407
Db      361 DNMAFLYRNISDVQTPTS-----TSSVSRSYNTNTIPLATSDVAVHGTGPDITL 412
Qy      408 PPT-SVMPVSPAPASTQYIGIHGTHYLLMPLYFIKRRKRVYFPADGEVA 460
Db      413 PPTIPQWNIYVPLPNNHSHVLOGTNTYLLPNYIFIKRRKRVYFPLTDEGVA 466

```

## RESULT 3

```

Q993Z4 PRELIMINARY; PRT; 473 AA.
ID AC Q993Z4
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)
DE Putative minor capsid protein L2.
GN L2.
OS Human papillomavirus type 82.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=129724;
RN [1]
RP SEQUENCE FROM N.A.
RA Terai M., Burk R.D.;
RT "Cervical HPV's in Evolution; Genomic Sequence of IS39/AE2, a Subtype
RT of Oncogenic HPV 82 (W13B).";
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF293961; AAK28455.1; -.
DR InterPro; IPR000784; Late_L2.
DR Pfam; PF00513; Late_protein_L2; 1.
SQ SEQUENCE 473 AA; 51047 MW; 3E31D287E0668FD3 CRC64;

```

Query Match 65.0%; Score 1563; DB 12; Length 473;  
 Best Local Similarity 62.8%; Pred. No. 8.3e-97;  
 Matches 297; Conservative 75; Mismatches 87; Indels 14; Gaps 7;

```

Qy      1 MWSHRAARRKRASTYDLYKTKQSGTSPDVYVKNVEGTTTLADKILQWSSLGIFLGIG 60
Db      1 MWSHRAARRKRASTYDLYKTKQSGTSPDVYVKNVEGTTTLADKILQWSSLGIFLGIG 60
Qy      61 TSGGTGRTGYIPLGGRNTIVDVSPAKRPVIEVPGTDPISIVTLIEDSSVITSGAPAP 119
Db      61 TSGGTGRTGYIPLGGRNTIVDVSPAKRPVIEVPGTDPISIVTLIEDSSVITSGAPAP 119
Qy      120 PFTGTSGFDITAGATTDAVIDITPSSVSISITNTFNPAFSDSIIIEVPGTGEVSGNV 179
Db      120 PFTGTSGFDITAGATTDAVIDITPSSVSISITNTFNPAFSDSIIIEVPGTGEVSGNV 179
Qy      180 VFGTSTGTHGEELIPLDTFASSGTCEPISSTPLPYARRVAGPRLYSRAVQVSVANPE 238
Db      181 IFSTPTSGTHGEELIPLDTFASSGTCEPISSTPLPYARRVAGPRLYSRAVQVSVANPE 240
Qy      239 PEFLTRPSLTITDNPFAFEPVDITLTF-EPNSNVPDSDFMIIIRLHRALTSRGTGVRFS 297
Db      241 PDLTRPSLTITDNPFAFEPVDITLTF-EPNSNVPDSDFMIIIRLHRALTSRGTGVRFS 300
Qy      298 RLQGRATMFTSRGTQIGARVHFHYHDISPASPEXIELQPLVSAETNGLEDIYAD----- 353
Db      301 RLQGRATMFTSRGTQIGARVHFHYHDISPASPEXIELQPLVSAETNGLEDIYADDEA 359
Qy      354 ---DIDPAMPVPSRPTTSSAVSYSPITSS--ASSYSNVTVPLTSSMDVAVYTGPDITLP 408

```

```

DB      | 360 ETCGCTPHTTTLKRSYSPLSTQPLSSSVSSSYANVTTPETSTYHPVHPGPPVLP 419
QY      | 409 PT-SWAPIVSPTAPASTQYIGHGHYIPLYFYIPKRRKRPVFEADGFA 460
        | : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      | 420 TSPYVFPFPHITSIDQHAIVLOGSDYIMPTIYLKRRKRRIFFADGFA 472

RESULT 4
056950
ID      | 056950      PRELIMINARY;      PRT;      519 AA.
AC      | 056950:
DT      | 01-JUN-1998 (TREMBLrel. 06, Created)
DT      | 01-JUN-1998 (TREMBLrel. 06, last sequence update)
DE      | 01-DEC-2001 (TREMBLrel. 19, last annotation update)
GN      | L2 protein.
OS      | Human papillomavirus type 77.
OC      | Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC      | Papillomavirus.
RX      | NCBI_TaxID=69986;
RN      | [1]
RP      | SEQUENCE FROM N.A.
RX      | MEDLINE=98118461; PubMed=9454709;
RA      | Dellus H., Saegling B., Bergmann K., Shamanin V., de Villiers E.M.;
RT      | "The genomes of three of four novel HPV types, defined by differences
RT      | of their L1 genes, show high conservation of the E7 gene and the
RT      | URR.";
RL      | Virology 240:359-365(1998).
DR      | EMBL: Y15175; CA75467.1; -
DR      | InterPro: IPR000784; Late_L2.
DR      | Pfam: PF00513; late-protein_L2; 1.
SQ      | SEQUENCE 519 AA; 56079 MW; 888F1407D9C114D0 CRC64;

Query Match
Best Local Similarity 63.7%; Score 1533; DB 12; Length 519;
Matches 302; Conservative 70; Mismatches 85; Indels 18; Gaps 10;

QY      | 1 MVSHRAARRKASVTDLYKTKQSGTSPDVVNKVGTTLADKILQWSSIGIFLAGIGIG 60
DB      | 47 MVAHRARRRRKRSATLKYCKAKGCPDPVLPKVGSTLADRIILQWGLGIVLGGIGIG 106
QY      | 61 TSSGTGRTGYIPLGSRNTVDVG-PTRPVYIEVPGTDPSTIVLIEDSSVYTGAPR 119
DB      | 107 TGTGTGGRTGYVPIGRPTVDVSVPTRPVYIEVPGTDPSTIVLIEDSSVYIDGASG 166
QY      | 120 PFTGSGFDITSAGTTTAVLDITPSSSVSISTNFPNPARSDSIIEVPTGSGN 179
DB      | 167 PFTGSGFEITSATTTTAVLDITPAGDTVYTSTNFPNLTETSEVLEVPQTGSGH 226
QY      | 180 VEVGPTSGTGHGEELIPLQTFASSSGTEEPISSTPLPTVRVAGPRLYSRAVOQSVANP 239
DB      | 227 LVTSTSGTGHGEELIPMDTFANSGTSEPISTPTPGVSRVAGPRLYGAMQVAVPDP 286
QY      | 240 EFLTRSSSLITDNPAFEPVDITLTFE---PNSNVDSDPMDIIRLHRPALISRGTVRF 296
DB      | 287 AFLSRSSSVTEPNPYDPGDETIIFERSPGTRVDPDLIDVRLHRPALTSRGTVRF 346
QY      | 297 SRLGATMETRSGTIGARVHFYHDISPIAPSPETIELOPIV-----SATDNGLDIYA 352
DB      | 347 SRKQGFMSKRTSGTIGARVHFYHDLSPTTHT-EDIELEPLLPDADSADSDS-LTDVYA 404
QY      | 353 --DDIDPAMPVPSRPTSSAVSTYSPGISASS--YSNVTVPLTSSMDVVPVYGPDTLP 408
DB      | 405 DVDDADVAFTNSGRNLTFSGGRASSSLPSALSTKGNVTIIPISVPDVALHMGPDVLP 464
QY      | 409 PYSWPIVSPTAPA-STOYIGHGHYIPLYFYIPK--RRVPYFEADGFA 460
DB      | 465 SSQWMPFV-PVLPADTTHYVYIDGNGFIAMPYTFVSRRKRRRLSFFADGFA 518

RESULT 5
08JUN96

```

```

ID      | 08JUN96      PRELIMINARY;      PRT;      467 AA.
AC      | 08JUN96:
DT      | 01-OCT-2002 (TREMBLrel. 22, Created)
DT      | 01-OCT-2002 (TREMBLrel. 22, last sequence update)
DE      | 01-MAR-2003 (TREMBLrel. 23, last annotation update)
GN      | Putative minor capsid protein L2.
OS      | Human papillomavirus type 90.
OC      | Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC      | Papillomavirus.
RX      | NCBI_TaxID=202251;
RN      | [1]
RP      | SEQUENCE FROM N.A.
RX      | MEDLINE=22079881; PubMed=12085327;
RA      | Terai M., Burk R.D.;
RT      | "Identification and Characterization of 3 Novel Genital Human
RT      | Papillomaviruses by Overlapping Polymerase Chain Reaction: candHPV89,
RT      | candHPV90, and candHPV91.";
RL      | J. Infect. Dis. 185:1794-1797(2002).
RN      | [2]
RP      | SEQUENCE FROM N.A.
RA      | Burk R.D., Terai M.;
RL      | Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR      | EMBL: AY057438; AAL14209.1; -
DR      | InterPro: IPR000784; Late_L2.
DR      | Pfam: PF00513; late-protein_L2; 1.
SQ      | SEQUENCE 467 AA; 49577 MW; 9BFC3FE24461D5C4 CRC64;

Query Match
Best Local Similarity 62.4%; Score 1507; DB 12; Length 467;
Matches 292; Conservative 70; Mismatches 94; Indels 12; Gaps 7;

QY      | 4 HRAARRKASVTDLYKTKQSGTSPDVVNKVGTTLADKILQWSSIGIFLAGIGIGS 63
DB      | 2 HRSKRRKRSATDLYKTKQAGTSPDVVNKVGTTLADKILQWSSIGIFLAGIGIGS 61
QY      | 64 GTGRTGYIPLGSRNTVDVG-PTRPVYIEVPGTDPSTIVLIEDSSVYTGAPRPT 122
DB      | 62 GTGRTGYIPIGTRPTVDVGPAPRPVYIEVPGASDPSTIVLIEDSSIIASGPHPT 121
QY      | 123 TGTSGFDITSAGTTTAVLDITPSSSVSISTNFPNPARSDSIIEVPTGSGN 182
DB      | 122 TGTGGEVNTASTTTPAVLDITPSSGNGVQSSSFTNPLETEPAIVEPPAGCVTGHVY 181
QY      | 183 GPTSGTGHGEELIPLQTFASSSGTEEPISSTPLPTVRVAGPRLYSRAVOQSVANP 242
DB      | 182 STPTASHGHEELIPMQTFANSGTSEPISTPLPTVRVAGPRLYSRAVOQSVANP 241
QY      | 243 TRPSSSLITDNPAFEPVDITLTFEPR-SNPVSDPMDIIRLHRPALTSRGTVRF 301
DB      | 242 SQEFLVTVYDNPVFEETLFEHPSIHQVDPDPLIDVLRHLPALARGTVRSRLQ 301
QY      | 302 RATMPTRSQIGARVHFYHDISPIAPSPETIELOPIVASTED--NGLPYIADDDPAM 359
DB      | 302 RATLKTSGSKRIATVHFYQDLSPVADLMQPLVSTPDPYDLSLVDIYADTASVS 361
QY      | 360 P--VPSRPTSSAVSTYSPGISASSYSNVTVPLTSSMDVVPVYGPDTLP 415
DB      | 362 HRLTPTREPTPLQASVYTNSSALSSAASNTYPLSTGLDIPVSGPDSALPSSHVMPV 421
QY      | 416 VSPTPAST--QYIGHGHYIPLYFYIPKRRKRPVFEADGFA 461
DB      | 422 --PPAPGVVGVSVLVNGSYIYLLPLGLLPKRRKRPPYFADGNVEA 467

RESULT 6
09JUN96
ID      | 09JUN96      PRELIMINARY;      PRT;      467 AA.
AC      | 09JUN96:
DT      | 01-OCT-2000 (TREMBLrel. 15, Created)
DT      | 01-OCT-2000 (TREMBLrel. 15, last sequence update)
DT      | 01-OCT-2002 (TREMBLrel. 22, last annotation update)
DE      | DNA, complete genome.

```

GN L2.  
 OS Human papillomavirus type 69.  
 OC Viruses; dsDNA viruses, no RNA stage: Papillomaviridae;  
 OC Papillomavirus.  
 NCBI\_TaxID=37121;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Matsukura T., Sata T.;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20087389; PubMed=10618284;  
 RA Kio N., Sata T., Sato Y., Sugase M., Matsukura T.;  
 RT "Molecular cloning and nucleotide sequence analysis of a novel human  
 papillomavirus (type 82) associated with vaginal intraepithelial  
 neoplasia.";  
 RL Clin. Diagn. Lab. Immunol. 7:91-95(2000).  
 DR EMBL; AB027020; BAA90733.1; -  
 DR InterPro: IPR000784; Late\_L2.  
 DR Pfam: PF00513; late\_protein\_L2; 1  
 SQ SEQUENCE 467 AA; 50054 MW; 78E32A50CECF9326 CRC64;

Query Match 62.2%; Score 1497; DB 12; Length 467;  
 Best Local Similarity 60.7%; Pred. No. 2,1e-92;  
 Matches 284; Conservative 81; Mismatches 93; Indels 10; Gaps 9;

QY 1 MWSHRAARRKASVTDLYKTKQSGTCDVYNNKVEGTTADKILQMSLGIPLGIG 60  
 DB 1 MVAAYASRRKRASADLTKCKACGTCPPDVIPKIEGSLADKILQMSGLGIFLGIG 60  
 QY 61 TGSAGGRTGYIPLGGRSN-TVVDYGPTRPPVIEPVPPTSTLTLEDSSVYTSAPR 119  
 DB 61 TGTGGRTGYIPLGGRSPSVVDIGPTRPILIEVGPTEPSTVTLVESSILQSGSP 120  
 QY 120 PFTGTSGDITTSAGTTPAVLDTTPSTSVSISTNTNPAFSDPSTIEVPQGEVSGN 179  
 DB 121 PNFSGDGEVYTTSTTPAVLDITPSPGVHTSTNQNPLIYEP-VDIQSGEALGH 179  
 QY 180 VFGVPTSGTGYEIPLOTFA--SGTEEPISSTPLTVRRVAGPRLYSRAVQVSYAN 238  
 DB 180 IFTSTGTGTHSYEIPMEVFAVSNMTSGSKPISSTPIGIRVAPRLYSKAYQOVKTD 239  
 QY 239 PEFTRPSLITTYDPAPEPVDITLTFEPRSNV--PDSPMDIIRLRALSRGTAFS 297  
 DB 240 PNFISKPSFTITPDNPAEPMDITLTFESADSHVADPDLIALHRALSRGTAFS 299  
 QY 298 RLGGATMFTSRGTQIGARVHYHDISPIASPEXIEIQPIVSATE-DNGLEDIYADDID 356  
 DB 300 RLGGATLTKTRSGKQIGAKVHYHDISPIATEIEIQPIITSEHSTPLEDYVA-DAD 358  
 QY 357 PAMPVPSRPTSSAVSTSPITSSASSY--NVTPLRSSMDVPRYTPDITLP-PTSVW 413  
 DB 359 PAMPFTPTPTPTPTPRSTSTSTSTSSAPLNVITPLSTSDIPYNDPIYAPVSSTW 418  
 QY 414 PIVSTPAASTQYIGIHGTHYLYPLLYFIPK-KRRKRYEFAAGFVA 460  
 DB 419 PYIPPPPTTMSHVAAGGNYYIMYIYLHKRRKRKRVCFSSGLAA 466

RESULT 7  
 Q9DHD4 PRELIMINARY; PRT; 472 AA.  
 AC Q9DHD4;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Putative L2 protein.  
 GN L2.  
 OS Human papillomavirus type 87.  
 OC Viruses; dsDNA viruses, no RNA stage: Papillomaviridae;  
 OC Papillomavirus.  
 NCBI\_TaxID=120381;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RA Menzo S., Monachetti A., Trozzi C., Delius H., Clementi M.;  
 RT "Complete nucleotide sequence of candidaPV87.";  
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ400628; CAC17717.2; -  
 DR InterPro: IPR000784; Late\_L2.  
 DR Pfam: PF00513; late\_protein\_L2; 1  
 SQ SEQUENCE 472 AA; 50340 MW; E740C9C3C5CD481B CRC64;

Query Match 61.2%; Score 1472.5; DB 12; Length 472;  
 Best Local Similarity 63.4%; Pred. No. 9,6e-91;  
 Matches 301; Conservative 61; Mismatches 86; Indels 27; Gaps 12;

QY 7 ARRRASVTDLYKTKQSGTCDVYNNKVEGTTADKILQMSLGIPLGIGTSGTG 66  
 DB 4 SRRKRASVTDLYKTKQSGTCDVYNNKVEGTTADKILQMSLGIPLGIGTSGTG 63  
 QY 67 GRTGYIPLGGRSNVVDYGPTRPPVIEPVPPTSTLTLEDSSVYTSAPRPTGT 125  
 DB 64 GRTGYIPLGRTPPVVDYGPTRPPVIEPVGADPSTVTLVESSVINSAGAPPNFTGT 123  
 QY 126 SGFDITSAGTTPAVLDTTPSTSVSISTNTNPAFSDPSTIEVPQGEVSGVFGTP 185  
 DB 124 GGFVYTTSTTPAVLDITPSSGVSYSTINPLTFEPSTIEVPQAGDIAGHYLSTA 183  
 QY 186 TSGTHGEYIPLQTF--SGTEEPISSTPLTVRRVAGPR--LYSRAVQVSYANPEFL 242  
 DB 184 TSGTHGEYIPLQTF--SGTEEPISSTPLTVRRVAGPR--LYSRAVQVSYANPEFL 243  
 QY 243 TPSSSLITTYDPAPEPVDITLTFEPRSNV--PDSPMDIIRLRALSRGTAFSRLQ 301  
 DB 244 SRPESVTFEHPMDPESTLTFEHPSLHAPDPDLITVLRALSRGTAFSRLQ 303  
 QY 302 RATFTSRGTQIGARVHYHDISPIASPEXIEIQPIVSAT--BDNGLEDIYADDI---- 355  
 DB 304 RASRTRSGKTIIGARVHYHDISPIA--DIEIQPIVSSTAPSDDTLXIDYADDITLSS 362  
 QY 356 --DPAMFV--PSRPTSSAVSTSPITSSASSYNTVPLTSSMDVPRYTPDITLP--T 406  
 DB 363 VLKRPSPSLTSSSTPTASTKVSAT--ILSSSHDNTVPL-SATDVPYLTGPDIDHSA 419  
 QY 407 LPPTSVMPPTPAASTQYIGIHGTHYLYPLLYFIPKRRKRYEFAAGFVA 461  
 DB 420 PSPPTPVVPSVTAIYS---YIQSDYLYLPLNTIIFPKRRKRKRVYFSGDFVA 471

RESULT 8  
 G9WNM5 PRELIMINARY; PRT; 471 AA.  
 AC G9WNM5;  
 DT 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE Minor capsid protein.  
 GN L2.  
 OS Human papillomavirus type 83.  
 OC Viruses; dsDNA viruses, no RNA stage: Papillomaviridae;  
 OC Papillomavirus.  
 NCBI\_TaxID=96240;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99335608; PubMed=10405368;  
 RA Brown D.R., McClellan T.L., Woods K., Fife K.H.;  
 RT "Nucleotide sequence and characterization of human papillomavirus type  
 83, a novel genital papillomavirus.";  
 RL Virology 260:165-172(1999).  
 GN [2]  
 RP SEQUENCE FROM N.A.  
 RA Brown D.R., McClellan T.L., Woods K., Fife K.H.;  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF151983; AAD38973.1; -  
 DR InterPro: IPR000784; Late\_L2.  
 DR Pfam: PF00513; late\_protein\_L2; 1.





```

Db      6 SRRKASATDLYRTCKASGTCPADVVPKVEGDTLADRIIKWASLGVFEGGLIGTSSGSG 65
QY      67 GRTGYIPLGGRSNFVVDVGP- RPPVIEPVGPTDPSIYTLIEDSSVYTSAGAPRPTFGT 125
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      66 GRTGYIPLGTRPPTVVDVGP- RPPVIEPVGPTDPSIYTLIEDSSVYTSAGAPRPTFGT 125
QY      126 SGFDITSAGTTPPAVLDTIPSSVSICTNFTNPAFSDSIIEVPQGEVSGNVFVGT 185
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      126 GGEFTVTSSTTPPAVLDTIPSSVSICTNFTNPAFSDSIIEVPQGEVSGNVFVGT 185
QY      186 TSGTGYEETPLQTEFA-SSGTGEEPISSPTLPTVRVAGPR--LYSRAVOQSVANPELT 242
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      186 TSGTGYEETPLQTEFA-SSGTGEEPISSPTLPTVRVAGPR--LYSRAVOQSVANPELT 245
QY      243 TRPSSLITVDNPAFEVDPTLTLEFRS--NVPSDFMDIIRLRPALTSRRCVRSRLG 300
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      246 SPPDSFVFPDNPVDP-DETLIEHPSLHAPDDPDLIVTLHRPALTRARSGVRSRLG 304
QY      301 QGATFTRSQTGICARVHFYHDSIPAPSEYIELOPLVSATNEDNG-----LFDIYAD 353
Db      305 QGASMTKRSKGIHIGARVHFYHDSIPAPSEYIELOPLVSATNEDNG-----LFDIYAD 362
QY      354 D--IDPAMPVPSRPTSSAVSTYSPYSS--ASSYSNVVPLTSSMDVPYTGPD-- 404
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      363 DTHLSIILOEPSPSLHSTTPVSSATVSAITSMVSTYDNTVPLSATEVPLTIGDDIDH 422
QY      405 -TLPPTSVMPVSPTPAPASTOY-IGIHGTHYIYLPYFYIPKRRKRVFFPADGVAA 461
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      423 AIAPSPTPLPVPV----PSTTPYAIYIGSDYLLPNYIFPFRKRRVHVSFSDGVAA 477

```

## RESULT 11

```

Q99FW8 PRELIMINARY; PRT; 474 AA.
ID      Q99FW8
AC      01-JUN-2001 (TREMblrel. 17, Created)
DT      01-JUN-2001 (TREMblrel. 17, Last sequence update)
DE      01-OCT-2001 (TREMblrel. 18, Last annotation update)
DE      Putative minor capsid protein L2.
GN      L2.
OS      Human papillomavirus type 84.
OC      Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC      Papillomavirus
OX      NCBI_TaxID=150546;
RN      [1]
RN      SEQUENCE FROM N.A.
RX      MEDLINE=2106735; PubMed=11145894;
RA      Teral M., Burk R.D.;
RT      *Complete Nucleotide Sequence and Analysis of a Novel Human
RT      Papillomavirus (HPV 84) Genome Cloned by an Overlapping PCR Method.*;
RL      Virology 279:109-115(2001).
RN      [2]
RN      SEQUENCE FROM N.A.
RA      Burk R.D., Teral M.;
RL      Submitted (Aug-2000) to the EMBL/genbank/DDbJ databases.
DR      EMBL: AF293960; AK09276.1; -
DR      InterPro: IPR000784; late_L2.
DR      Pfam: PF00513; late_protein_L2.
SQ      SEQUENCE 474 AA; 50608 MW; 04BD0ECAED298551 CRC64;

```

Query Match 59.4%; Score 1429.5; DB 12; Length 474;

Best Local Similarity 62.6%; Pred. No. 7.3e-88;

Matches 293; Conservative 55; Mismatches 105; Indels 15; Gaps 9;

```

QY      8 RKRASVYDLYTKTQSGTCEPDVYNNKYEGTTLADKIIQMSLGIIFLGIGTGSGTGG 67
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      7 RKRASVYDLYTKTCAIGTRPADVLPKVEGDTWADRFLKWSLGVFEGGLIGTGSGTGG 66
QY      68 RTGYIPLGGRSNFVVDVGP- RPPVIEPVGPTDPSIYTLIEDSSVYTSAGAPRPTFGT 126
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      67 RTGYIPLGTRPPTVVDVGP- RPPVIEPVGPTDPSIYTLIEDSSVYTSAGAPRPTFGT 126
QY      127 GPDITSAGTTPPAVLDTIPSSVSICTNFTNPAFSDSIIEVPQGEVSGNVFVGT 186

```

```

Db      127 GFEVTTSSITTPPAVLDTIPSSVSICTNFTNPAFSDSIIEVPQGEVSGNVFVGT 186
QY      187 SGTGHEEIPLOTEFA-SSGTGEEPISSPTLPTVRVAGPR--LYSRAVOQSVANPELT 243
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      187 SGTGHEEIPLOTEFA-SSGTGEEPISSPTLPTVRVAGPR--LYSRAVOQSVANPELT 246
QY      244 RPSLITVDNPAFEVDPTLTLEFRS--NVPSDFMDIIRLRPALTSRRCVRSRLG 302
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      247 RPSLITVDNPAFEVDPTLTLEFRS--NVPSDFMDIIRLRPALTSRRCVRSRLG 306
QY      303 ATMTFRSGTQICARVHFYHDSIPAPSEYIELOPLVSATNEDNG-----LFDIYAD- IDPA 358
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      307 ASMRTRSKGIHIGARVHFYHDSIPAPSEYIELOPLVSATNEDNG-----LFDIYAD 365
QY      359 MPVPS-----RPTSSAVSTYSPYSS--ASSYSNVVPLTSSMDVPYTGPDITLPTSVMP 414
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      366 LRPVSALRKASPPASADLSITASTYDNTVPLSGVDVPTTTPDDIDHSAAPSAP 425
QY      415 IVSPTAPASTOY-IGIHGTHYIYLPYFYIPKRRKRVFFPADGVAA 461
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      426 PVPVTPSTTPYAIYIGSDYLLPNYIFPFRKRRVHVSFSDGVAA 473

```

## RESULT 12

```

Q82002 PRELIMINARY; PRT; 466 AA.
ID      Q82002
AC      082002;
DT      01-NOV-1996 (TREMblrel. 01, Created)
DT      01-NOV-1996 (TREMblrel. 01, Last sequence update)
DE      01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE      Late gene.
GN      L2.
OS      Human papillomavirus type 72.
OC      Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC      Papillomavirus
OX      NCBI_TaxID=51032;
RN      [1]
RN      SEQUENCE FROM N.A.
RX      MEDLINE=96213783; PubMed=8635859;
RA      Voelter C., He Y., Dellus H., Roy-Burman A., Greenspan J.S.,
RT      *Novel HPV types present in oral papillomatous lesions from patients
RT      with HIV infection.*;
RL      Int. J. Cancer 66:453-456(1996).
DR      EMBL: X94164; CA63878.1; -
DR      InterPro: IPR000784; late_L2.
DR      Pfam: PF00513; late_protein_L2.
SQ      SEQUENCE 466 AA; 49709 MW; B3D06F43227BDAA0 CRC64;

```

Query Match 58.9%; Score 1417.5; DB 12; Length 466;

Best Local Similarity 61.1%; Pred. No. 4.6e-87;

Matches 286; Conservative 66; Mismatches 93; Indels 23; Gaps 10;

```

QY      8 RKRASVYDLYTKTQSGTCEPDVYNNKYEGTTLADKIIQMSLGIIFLGIGTGSGTGG 67
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      7 RKRASVYDLYTKCAIGTRPADVLPKVEGDTWADRFLKWSLGVFEGGLIGTGSGTGG 66
QY      68 RTGYIPLGGRSNFVVDVGP- RPPVIEPVGPTDPSIYTLIEDSSVYTSAGAPRPTFGT 126
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      67 RTGYIPLGTRPPTVVDVGP- RPPVIEPVGPTDPSIYTLIEDSSVYTSAGAPRPTFGT 126
QY      127 GPDITSAGTTPPAVLDTIPSSVSICTNFTNPAFSDSIIEVPQGEVSGNVFVGT 186
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      187 SGTGHEEIPLOTEFA-SSGTGEEPISSPTLPTVRVAGPR--LYSRAVOQSVANPELT 243
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      187 SGTGHEEIPLOTEFA-SSGTGEEPISSPTLPTVRVAGPR--LYSRAVOQSVANPELT 246
QY      244 RPSLITVDNPAFEVDPTLTLEFRS--NVPSDFMDIIRLRPALTSRRCVRSRLG 301
        |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      247 RPSLITVDNPAFEVDPTLTLEFRS--NVPSDFMDIIRLRPALTSRRCVRSRLG 305

```



```
RESULT 15
ID 081016 PRELIMINARY; PRT; 466 AA.
AC 081016;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE L2 capsid protein.
OS Human papillomavirus type 31.
OC Viruses; dsDNA viruses, no RNA stage; Papillomaviridae;
OC Papillomavirus.
OX NCBI_TaxID=10585;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MS2;
RX MEDLINE=96130211; PubMed=8553573;
RA Icenogle J.P., Clancy K.A., Lin S.Y.;
RT "Sequence variation in the capsid protein genes of human
RL papillomavirus type 16 and type 31.";
RL Virology 214:664-669(1995).
DR EMBL: U37410; AAA92893.1; -.
DR InterPro: IPR000784; Late_L2.
DR Pfam: PF00513; late-protein_L2; 1.
SQ SEQUENCE 466 AA; 50040 MW; 125438BD13A83412 CRC64;

Query Match 48.3%; Score 1160.5; DB 12; length 466;
Best Local Similarity 50.9%; Pred. No. 7,66-70;
Matches 247; Conservative 78; Mismatches 117; Indels 43; Gaps 15;

QY 1 MYSRAARR-KRASVDLYKCTCKQSGCPDVKVNEGTTLADKILQWSSLGIFLGIGI 59
   | | : : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 MRSKRSTKRTKRASATOLYQTKACAGCPDSDVIRKIEHTTIADQILRYSMGVFEFGIGI 60
-QY 60 GTSGGTGGRGYIILGGSNTYVDG--PTRBPVYIEPYGTPDPSPVTLIEDSSVYTSAP 118
   | : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 GSGGTGGRGYIPLSTPSTVSEASIPRPVSIDPVGPLDPSTVSLVESGIIDVGAP 120
QY 119 R-PTFGTSGFDITSAQTTPPAVIDITPSSVSISTNFTNPAPFSDPSIIEVPTGCV 176
   | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 APIRPRPTSGFDIATADTAPALDVT-----SVST--HENPTDPSTVLQPTPAET 172
QY 177 SGANFVGTPSGTGYEIEPIQITFASGSGEGEPISSPTPLTVRRYACPRLYSKAYQOVS 236
   | : : : : | | | | | | | | | | | | | | | | | | | | | |
Db 173 SGHLLSSSSISSTHNYEIIPMDTFIVSTNNENITSTPIPGVRRPARGLYSKATQYKV 232
QY 237 ANPEFLRPSSLITYNDPAPEPV--DTTLTEPRSN--VPDSDFMDIIRLHRPALTSRG 292
   : | | : | | | | | | | | | | | | | | | | | | | | | | |
Db 233 IDPFTLSAPKOLITYENPAITVNAESLFSNTSHMKAPDPDFLDITLHRPALTSRRN 292
QY 293 TVRPSRLGORATMFTRSQTGIGARVHFYHDISPIAPSPXYTELQPL-----VSATEDNGL 347
   | | | | | | : | | | | | | | | | | | | | | | | | | | |
Db 293 TVRSRLGNKQTLTRSGATIGARVHYIIDSSINPAGESIEMQPLGASATTTSTINDGL 352
QY 348 FDIYADIDPAMPVPS-----RPTT-----SSAVSTYSPTISSASYSNTVPLTSSMDVPV 399
   : | | | | | : : | | : | | | | | | | | | | | | | |
Db 353 YDIYA-DTDFVVDPPATHNVPSTALQSTSAVSAYVP-----NTTVPLSTGFDIPI 403
QY 400 YTGPDITL--PPTSVWIVSPATAPASQY-IGIGHTHYLLPLVYFIKRRKRVYPFAD 456
   : | | | : | | | | | | | | | | | | | | | | | | | |
Db 404 FSGPDVPIEHAFTOVFP--PLAFTTPOVSLFVDGDFYLLHPSYMLKRRKRVSYPFTD 461
QY 457 GFVAA 461
   |||
Db 462 VSVAA 466
```

Search completed: October 17, 2003, 11:02:08  
Job time : 66.9112 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2003, 11:02:14 : Search time 47.6622 Seconds  
(without alignments)  
1743.646 Million cell updates/sec

Title: US-08-913-644-2

Perfect score: 2733  
Sequence: 1 MALMRSDNTVYLPPEVSAR.....APSATSTSSKPAKRVRRARK 507

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 609560 seqs, 163917102 residues

Total number of hits satisfying chosen parameters: 609560

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

Published\_Applications\_AA.\*  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2705	99.0	568	11	US-09-991-053-3
2	2705	99.0	568	15	US-10-301-260A-3
3	1831	67.0	505	10	US-09-162-904A-4
4	1826	66.8	505	10	US-09-824-017-2
5	1826	66.8	505	11	US-09-986-118A-2
6	1825	66.8	505	10	US-09-162-904A-3
7	1820.5	66.6	502	12	US-10-224-999A-3469
8	1801	65.9	531	11	US-09-991-053-1
9	1801	65.9	531	15	US-10-301-260A-1
10	1324.5	48.5	508	9	US-09-134-377-1
11	1298	47.5	508	10	US-09-822-662-1
12	1298	47.5	503	9	US-09-134-377-2
13	1298	47.5	503	10	US-09-822-662-2
14	1283	46.9	495	12	US-10-305-765-2
15	1283	46.9	495	12	US-10-305-765-4

16	1283	46.9	495	12	US-10-305-633-2	Sequence 2, Appl1
17	1283	46.9	495	12	US-10-305-633-4	Sequence 4, Appl1
18	604	22.1	132	14	US-10-056-359-29	Sequence 29, Appl1
19	604	22.1	132	14	US-10-056-360-29	Sequence 29, Appl1
20	551.5	20.2	151	12	US-10-224-999A-3467	Sequence 3467, Ap
21	374	13.7	144	14	US-10-056-359-5	Sequence 5, Appl1
22	374	13.7	144	14	US-10-056-360-5	Sequence 5, Appl1
23	363	13.3	136	14	US-10-056-359-8	Sequence 8, Appl1
24	363	13.3	136	14	US-10-056-360-8	Sequence 8, Appl1
25	359.5	13.2	139	14	US-10-056-359-2	Sequence 2, Appl1
26	359.5	13.2	139	14	US-10-056-360-2	Sequence 2, Appl1
27	353.5	12.9	145	14	US-10-056-359-11	Sequence 11, Appl1
28	353.5	12.9	145	14	US-10-056-360-11	Sequence 11, Appl1
29	353	12.9	138	14	US-10-056-359-23	Sequence 23, Appl1
30	353	12.9	138	14	US-10-056-360-23	Sequence 23, Appl1
31	352	12.9	138	14	US-10-056-359-14	Sequence 14, Appl1
32	352	12.9	138	14	US-10-056-360-14	Sequence 14, Appl1
33	351	12.8	142	14	US-10-056-359-20	Sequence 20, Appl1
34	351	12.8	142	14	US-10-056-360-20	Sequence 20, Appl1
35	333.5	12.2	137	14	US-10-056-359-17	Sequence 17, Appl1
36	333.5	12.2	137	14	US-10-056-360-17	Sequence 17, Appl1
37	331.5	12.1	141	14	US-10-056-359-26	Sequence 26, Appl1
38	331.5	12.1	141	14	US-10-056-360-26	Sequence 26, Appl1
39	109	4.0	967	12	US-10-032-189-38	Sequence 38, Appl1
40	101.5	3.7	5701	9	US-09-864-761-37319	Sequence 37319, A
41	100.5	3.7	26926	10	US-09-759-508B-2	Sequence 2, Appl1
42	94.5	3.5	710	15	US-10-296-770-4	Sequence 4, Appl1
43	94.5	3.5	777	10	US-09-765-111A-2	Sequence 2, Appl1
44	90.5	3.3	788	14	US-10-003-152-6	Sequence 6, Appl1
45	90.5	3.3	788	15	US-10-002-050-6	Sequence 6, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-991-053-3  
; Sequence 3, Application US/09991053  
; Publication No. US20030003532A1  
; GENERAL INFORMATION:  
; APPLICANT: Harrison, Stephen  
; APPLICANT: Chen, Xiaojiang  
; TITLE OF INVENTION: Vaccine compositions and methods for human papilloma  
; TITLE OF INVENTION: virus  
; FILE REFERENCE: Harvard/Harrison 12687/1120  
; CURRENT APPLICATION NUMBER: US/09/991,053  
; CURRENT FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: 09/520,781  
; PRIOR FILING DATE: 2000-03-18  
; PRIOR APPLICATION NUMBER: 60/125208  
; PRIOR FILING DATE: 1999-03-18  
; PRIOR APPLICATION NUMBER: 60/148544  
; PRIOR FILING DATE: 1999-08-12  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 568  
; TYPE: PRT  
; ORGANISM: Human papillomavirus type 18  
US-09-991-053-3

Query Match 99.0%; Score 2705; DB 11; Length 568;  
Best Local Similarity 99.0%; Pred. No. 7.6e-271;  
Matches 502; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY	1	MALMRSDNTVYLPPEVSARVNTDDYVTRTSIFYHAGSSRLLTGNGPFRVAGGAKNO	60
DB	62	MALMRSDNTVYLPPEVSARVNTDDYVTPTSIFYHAGSSRLLTGNGPFRVAGGAKNO	121
QY	61	DIFKVSAYQVRVRYVQLPDNRKFGLPDNSLYNETORLWACAGVEIGRGQPIGVGISGH	120
DB	122	DIFKVSAYQVRVRYVQLPDNRKFGLPDTSIFYNETORLWACAGVEIGRGQPIGVGISGH	181

Qy	121	PFYFNKDDIHESSIAATSANSEEDVRNANVSDYQDOLCIIIGCAPAIGHEMAGTACKRPL	180
	182	PFYFNKDDIHESSIAATSANSEEDVRNANVSDYQDOLCIIIGCAPAIGHEMAGTACKRPL	241
Db	181	SQGDCCPPELEIKNTVLIEDGDMDVTGYGAMDFTLQDCKEVPBLDIOCSICKKPYDLQMSAD	240
Db	242	SQGDCCPPELEIKNTVLIEDGDMDVTGYGAMDFTLQDCKEVPBLDIOCSICKKPYDLQMSAD	301
Qy	241	PYGDSMFECFLRRQOLFARHFENRAGTMGDYVQSLYIKGTGKRASPGSCVYSPSPSGSIV	300
Db	302	PYDPSMFECFLRRQOLFARHFENRAGTMGDYVQSLYIKGTGKRASPGSCVYSPSPSGSIV	361
Qy	301	TTSOSOLEFNKPYMLHKRQGNNGICQHNOLFVYVDTTRSTNLITTCASQSPGQDYATK	360
Db	362	TTSOSOLEFNKPYMLHKRQGNNGICQHNOLFVYVDTTRSTNLITTCASQSPGQDYATK	421
Qy	361	FKQYSRHVEEYDLOFTFOLCTITLLADWASYIHSNMSSILLEDWNGVPPPTTSLVDYTR	420
Db	422	FKQYSRHVEEYDLOFTFOLCTITLLADWASYIHSNMSSILLEDWNGVPPPTTSLVDYTR	481
Qy	421	FVQSVAVITCQKDAAPAKENKOPYDKLKFNVNDLKEFSLDDQYPLGKRFVQAGLRKRP	480
Db	482	FVQSVAVITCQKDAAPAKENKOPYDKLKFNVNDLKEFSLDDQYPLGKRFVQAGLRKRP	541
Qy	481	IGPRKRSAPSATSSKPAKRRVVRAR	507
Db	542	IGPRKRSAPSATSSKPAKRRVVRAR	568

```

RESULT 2
US-10-301-260A-3
Sequence 3, Application US/10301260A
Publication No. US20030118609A1
GENERAL INFORMATION:
APPLICANT: Harrison, Stephen
APPLICANT: Chen, Xiaojiang
TITLE OF INVENTION: Vaccine compositions and methods for human papilloma
FILE REFERENCE: Harvard/Harrison 12687/1123
CURRENT APPLICATION NUMBER: US/10/301,260A
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: 09/520,822
PRIOR FILING DATE: 2000-03-18
PRIOR APPLICATION NUMBER: 60/125208
PRIOR FILING DATE: 1999-03-18
PRIOR APPLICATION NUMBER: 60/148544
PRIOR FILING DATE: 1999-08-12
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 568
TYPE: PRF
ORGANISM: Human papillomavirus type 18
US-10-301-260A-3

```

	Query Match	99.0%	Score 2705:	DB 15:	Length 568:
	Best Local Similarity	99.0%	Pred. No. 7,66-271:		
	Matches	502:	Conservative	1:	Mismatches 4: Indels 0: Gaps 0:
QY	1	MALMRPSDNYYLPPEPSAVRVNTDDYYTRISIFYHAGSSRLTYGNFYFRYPAGCGNKQ	60		
Db	62	MALMRPSDNYYLPPEPSAVRVNTDDYYTPSIFYHAGSSRLTYGNFYFRYPAGCGNKQ	121		
QY	61	DIPKVSAYQYKVERFVOLDPDKFGLPDMISINPETORLVMACAGYEIGRGPVIGLSGH	120		
Db	122	DIPKVSAYQYKVERFVOLDPDKFGLPDMISINPETORLVMACAGYEIRGQPLVGLSGH	181		
QY	121	PFTYKKLDDTSSHAATSVNSEDVDGNSVDVKQKQOLCLGCAPAIGEHMAGTACKSPL	180		
Db	182	PFTYKKLDDTSSHAATSVNSEDVDGNSVDVKQKQOLCLGCAPAIGEHMAGTACKSPL	241		
QY	181	SQGDGPLELKNVILEDGWVDTGAGADESTLDQFKCEVPDLDTQSICTKPYDTLOMSAD	240		

Db	24	SQGDCEPBLEKNTLVLEDGSMVDTGVGAMDFSLDQTKCEVPLDIOQISCTKPYDILQMSAD	301
Qy	241	PYGDSMFECLLRQQLARHFEFNRAGTMDPTVQSLYIKGTGMRASPGSCVYSPSPSGSIV	300
Db	302	PYDSDMFECLLRQQLARHFEFNRAGTMDPTVQSLYIKGTGMRASPGSCVYSPSPSGSIV	361
Qy	301	TSDSOLEFNKPYWLHKAGQANNICWHDLEFVIVDTTRSTNLTICASTOSFVPCQYATK	360
Db	362	TSDSOLEFNKPYWLHKAGQANNICWHDLEFVIVDTTRSTNLTICASTOSFVPCQYATK	421
Qy	361	EKQSYHVEYDQLEFQCLCTITLADWASYIHSNNSLTLEDMNGVPPPTTSLVDTYR	420
Db	422	EKQSYHVEYDQLEFQCLCTITLADWASYIHSNNSLTLEDMNGVPPPTTSLVDTYR	481
Qy	421	FVQSVATTCQDAAPAEKNDPYDKLTFWNVDLKEKFSLDLDQYPLGRKFLVQAGILRRPT	480
Db	482	FVQSVATTCQDAAPAEKNDPYDKLTFWNVDLKEKFSLDLDQYPLGRKFLVQAGILRRPT	541
Qy	481	IGPRKRSAPGATSSSKPAKRVVPRARK	507
Db	542	IGPRKRSAPGATSSSKPAKRVVPRARK	568

```

RESULT 3
US-09-162-904A-4
Sequence 4, Application US/09162904A
Patent No. US2002016872A1
GENERAL INFORMATION:
APPLICANT: Durst, Mathias
APPLICANT: Gissmann, Lutz
TITLE OF INVENTION: A DNA SEQUENCE ENCODING A PAPILLOMAVIRUS L1 PROTEIN CAPABLE
TITLE OF INVENTION: OF EFFICIENTLY FORMING VIRUS-LIKE PARTICLES
FILE REFERENCE: 8484-068-999
CURRENT APPLICATION NUMBER: US/09/162,904A
CURRENT FILING DATE: 1998-09-29
PRIOR APPLICATION NUMBER: 08/884,168
PRIOR FILING DATE: 1997-06-27
PRIOR APPLICATION NUMBER: 08/641,570
PRIOR FILING DATE: 1996-05-01
PRIOR APPLICATION NUMBER: 08/902,528
PRIOR FILING DATE: 1993-07-16
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn version 3.1
SEQ ID NO 4
LENGTH: 505
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Amino acid sequence of L1
US-09-162-904A-4

```

Query Match	67.0%	Score 1831	DB 10	Length 505
Best Local Similarity	66.5%	Pred. No. 26-180		
Matches	334	Conservative	68	Mismatches 96; Indels 4; Gaps 3;
Qy	1	MALRRPSDNVYLLPPPSVARYVNDIDYTRISIEYHAGSSRLITVGNPFYRVAPAGGNKQ	60	
Db	1	MSLMPSATVYLLPPVPVSKVYSTEDEYAKRTIYHAGTSRLLAGVHPPIKPPNNKRI	60	
Qy	61	DIPKSAQYRYRFRQLDPNKRFGLDPSIYNPEFQRLVMACAGVEIGGQPLGVLISGH	120	
Db	61	LVPKVSGLQYVFRHLHPDPNKRFGFPDPSFYNPDPQRLVMACVGVGVGGQPLGVGISGH	120	
Qy	121	PFYKKLDDTESSHAATSVSESDVDRDVSVDKIQOTLCICGAPAIGEHMAKTSKSRPL	180	
Db	121	PLTKKLDDTEENASAAVANAAGVDNRRCISMDYKQOTLCICGKRPPIGEHMGKSPCTNNVAV	180	
Qy	181	SGQGCPLPELKNVYLEDGDMDVTGCGAMDEFSTLOOTKCEVPLDIOSTICKYDYLOMSAD	240	
Db	181	NPQGCPLPELKNVYIQDGDMDVTGCGAMDEFSTLOANKSVPLDICTSICKYDYIIMWSE	240	
Qy	241	PYGDSEPFCLRRDLEAFHFVNNRACMGIDYVQSLYIKGTGNRASPGSCVSPSPGSLIV	300	



TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-986-118A-2

Query Match 66.8%; Score 1826; DB 11; Length 505;  
Best Local Similarity 66.3%, Pred. No. 6,7e-180;  
Matches 333; Conservative 68; Mismatches 97; Indels 4; Gaps 3;

OY 1 MALRRPSDNYVLPPEPPSARAVYNDDYTRKISIFTHAGSSMLLVGNPFVNPAGGNGKQ 60  
Db 1 MSLLPSPETATVLPVPVSKVSDEYVARNTIYHAGTSILLVAGHFYFKRPNNKI 60  
OY 61 DIPVSAVQYRVFVFOVLDPDAKFGLPDNSIYNPETORLVMACAGEIGRGPPLGVLSGH 120  
Db 61 LVPRVSGIQYVFRIHLHPDPKKFGFPDISEFNPTQRILVMACGEVEVGROGPLGVLSGH 120  
OY 121 PFYVKLDDTESSHAATSINSEVDVNDNVSDIKQTQLCILGCAPAI GEHWAKGTACKSRPL 180  
Db 121 PLTKRLDDTEVASAYANAGVNRECI SMDKYOTLCIGCKPIGEHWKGSPCTNVAV 180  
OY 181 SOGGCPPLELNVTLEGDWDVDTGYGAMDFTSLDTDFCEVFLDIQSICXPDLWMSAD 240  
Db 181 NPGCCPPELINTYTIQDGWDVDTFGAMDFITLDANKSEVPLDICTSCIKPIDIKMWE 240  
OY 241 PYGDSMEFCFLREQLFAHFENNRAGTMDVPOSLYLKGTGMRASPGSCVYSFPSGSIV 300  
Db 241 PYGSLFEFYLRREOMFYHLFNRAAGVENPDDL IYIGSGSTANIASSNFPPSGSMV 300  
OY 301 TSDSOLFKNRPWLKAOGHNNGICMHNOLEFVVVDTRSTULITCASQSPVPOQYDATK 360  
Db 301 TSDAQIRFKRPWLQRAQHNGICMGNDLFVTVDITRTSYMSICAAL-STSETTYKNTN 359  
OY 361 FKQYSRHVEYDLOFIQLCTITLTADVMSYHSNMSSILEDMMNFEGVPPPTSLVDYTR 420  
Db 360 FKEYLHRHEEYDLOFIQLCKITTLADVMYTIHSMNSTILEDMMNFGLPPPGGLEPTYR 419  
OY 421 FVQSVAITCQKDAPAENKDKDYDLKTFMNVDLKEKFSLDLQYPLGRFTLVQAQIRKP- 479  
Db 420 FVTSQALACQKHPTPAEDPLKRYTFEMVNLKERFSADLDQFLGRKFTLLQAGLAKPK 479  
OY 480 -TIGPRKSAPSATTSSKPKARR 500  
Db 480 FTLGKRK-ATPTTSTSTTAKR 500

RESULT 6  
US-09-162-904A-3  
Sequence 3, Application US/09162904A  
Patent No. US20020168372A1  
GENERAL INFORMATION:  
APPLICANT: Durest, Matthias  
APPLICANT: Gissmann, Lutz  
TITLE OF INVENTION: A DNA SEQUENCE ENCODING A PAPILLOMAVIRUS LI PROTEIN CAPABLE  
FILE REFERENCE: 8484-068-999  
CURRENT APPLICATION NUMBER: US/09/162,904A  
CURRENT FILING DATE: 1998-09-29  
PRIOR APPLICATION NUMBER: 08/884,168  
PRIOR FILING DATE: 1997-06-27  
PRIOR APPLICATION NUMBER: 08/641,570  
PRIOR FILING DATE: 1996-05-01  
PRIOR APPLICATION NUMBER: 08/902,528  
PRIOR FILING DATE: 1993-07-16  
NUMBER OF SEQ ID NOS: 6  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 505.  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: HPV16 clone  
OTHER INFORMATION: Description of Artificial Sequence: Amino acid sequence of LI gene

[illegible]



```

QY 61 DLPKSAIYQYRFRVQLPDPNKGFLPDNSIYNFETQRLVWACAGVEIGRGQPLGVLGSH 120
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 AYPKVSGLQYRFRVRLPDPNKGFLPDNSIYNFETQRLVWACAGVEIGRGQPLGVLGSH 120
QY 121 PFYNTKLDDESSHAATSNVSEYRDVNVSDYKOTQCLIGCAPAIGEHNAKGTACKSRPL 180
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 PLINKLDDTENASAVYAMAGVDRNRECI SMYKOTQCLIGCKPPIEBHMGKSPCTNVAV 206
QY 181 SOGDCEPPELKNVTLEDGMDVDTGYGAMPESTLQDKCEVPLDICSICKPYDYLQMSAD 240
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 KAGECPPELKNVTLEDGMDVDTGYGAMPESTLQDKCEVPLDICSICKPYDYLQMSAD 240
QY 241 PFGDSMFECFLRRQLEFARHFNAGTMDTPVPSLYIKGTGMARSPGSCVSPSPGSIY 300
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 PYGDSMFECFLRRQLEFARHFNAGTMDTPVPSLYIKGTGMARSPGSCVSPSPGSIY 300
QY 301 TSDSOLEFNRPYMLHQAQGNHNGICWGNOLFVTVVDTTRSTNMLTICASTOSPVGQYDARK 360
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 TSDSOLEFNRPYMLHQAQGNHNGICWGNOLFVTVVDTTRSTNMLTICASTOSPVGQYDARK 360
QY 361 FKQYSHREBYDQLEFQOLCTITLTDVNSYIHSNNSILEDNMFVGPPEPTSLVDYR 420
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 FKQYSHREBYDQLEFQOLCTITLTDVNSYIHSNNSILEDNMFVGPPEPTSLVDYR 420
QY 421 FVQSAVITQCKDAAPAEKNDPYDKLFMNVDLKEKFSLDLDQPLGRKFLVQAGLRKP 479
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 421 FVQSAVITQCKDAAPAEKNDPYDKLFMNVDLKEKFSLDLDQPLGRKFLVQAGLRKP 479
QY 446 FV-QAIAQCKHTPPAPKEDDPLKKTTFEVLNKEKFSADLDQPLGRKFLVQAGLRKP 504
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 446 FV-QAIAQCKHTPPAPKEDDPLKKTTFEVLNKEKFSADLDQPLGRKFLVQAGLRKP 504
QY 480 --TIGPKRSAPSATSSKPAKR 500
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 480 --TIGPKRSAPSATSSKPAKR 500
QY 481 IGPKRKSAPSATSSKPAKR 505
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 481 IGPKRKSAPSATSSKPAKR 505
QY 478 FRLGKRAAPASISKSTRRKVK 502
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 478 FRLGKRAAPASISKSTRRKVK 502

```

```

RESULT 8
US-09-991-053-1
; Sequence 1, Application US/09991053
; Publication No. US20030003532A1
; GENERAL INFORMATION:
; APPLICANT: Harrison, Stephen
; APPLICANT: Chen, Xiaojiang
; TITLE OF INVENTION: Vaccine compositions and methods for human papilloma
; TITLE OF INVENTION: virus
; FILE REFERENCE: Harvard/Harrison 12687/1120
; CURRENT APPLICATION NUMBER: US/09/991,053
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: 09/520,781
; PRIOR FILING DATE: 2000-03-18
; PRIOR APPLICATION NUMBER: 60/125208
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/148544
; PRIOR FILING DATE: 1999-08-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO: 1
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-09-991-053-1

```

```

Query Match 65.9%; Score 1801; DB 11; Length 531;
Best Local Similarity 66.0%; Pred. No. 2,8e-177;
Matches 332; Conservative 69; Mismatches 96; Indels 6; Gaps 5;
QY 1 MALMRSDMTVYLPPEPSAVRVNTDDYTRTSIFYHAGSSRLTGVNPFYRVVAGGANKQ 60
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 27 MSIMLPSEAVTVLPPEPSAVRVNTDDYTRTSIFYHAGSSRLTGVNPFYRVVAGGANKQ 60
QY 61 DIPKVSAYQYRFRVQLPDPNKGFLPDNSIYNFETQRLVWACAGVEIGRGQPLGVLGSH 120
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 DIPKVSAYQYRFRVQLPDPNKGFLPDNSIYNFETQRLVWACAGVEIGRGQPLGVLGSH 120
QY 87 LVPKVSGLQYRFRVRLPDPNKGFLPDNSIYNFETQRLVWACAGVEIGRGQPLGVLGSH 146
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 87 LVPKVSGLQYRFRVRLPDPNKGFLPDNSIYNFETQRLVWACAGVEIGRGQPLGVLGSH 146
QY 121 PFYNTKLDDESSHAATSNVSEYRDVNVSDYKOTQCLIGCAPAIGEHNAKGTACKSRPL 180
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 PFYNTKLDDESSHAATSNVSEYRDVNVSDYKOTQCLIGCAPAIGEHNAKGTACKSRPL 180

```

```

Db 147 PLINKLDDTENASAVYAMAGVDRNRECI SMYKOTQCLIGCKPPIEBHMGKSPCTNVAV 206
QY 181 SOGDCEPPELKNVTLEDGMDVDTGYGAMPESTLQDKCEVPLDICSICKPYDYLQMSAD 240
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 207 NPEDCEPPELKNVTLEDGMDVDTGYGAMPESTLQDKCEVPLDICSICKPYDYLQMSAD 266
QY 241 PFGDSMFECFLRRQLEFARHFNAGTMDTPVPSLYIKGTGMARSPGSCVSPSPGSIY 300
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 PYGDSMFECFLRRQLEFARHFNAGTMDTPVPSLYIKGTGMARSPGSCVSPSPGSIY 300
QY 301 TSDSOLEFNRPYMLHQAQGNHNGICWGNOLFVTVVDTTRSTNMLTICASTOSPVGQYDARK 360
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 TSDSOLEFNRPYMLHQAQGNHNGICWGNOLFVTVVDTTRSTNMLTICASTOSPVGQYDARK 360
QY 327 TSDSOLEFNRPYMLHQAQGNHNGICWGNOLFVTVVDTTRSTNMLTICASTOSPVGQYDARK 385
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 327 TSDSOLEFNRPYMLHQAQGNHNGICWGNOLFVTVVDTTRSTNMLTICASTOSPVGQYDARK 385
QY 361 FKQYSHREBYDQLEFQOLCTITLTDVNSYIHSNNSILEDNMFVGPPEPTSLVDYR 420
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 FKQYSHREBYDQLEFQOLCTITLTDVNSYIHSNNSILEDNMFVGPPEPTSLVDYR 420
QY 421 FVQSAVITQCKDAAPAEKNDPYDKLFMNVDLKEKFSLDLDQPLGRKFLVQAGLRKP 479
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 421 FVQSAVITQCKDAAPAEKNDPYDKLFMNVDLKEKFSLDLDQPLGRKFLVQAGLRKP 479
QY 446 FV-QAIAQCKHTPPAPKEDDPLKKTTFEVLNKEKFSADLDQPLGRKFLVQAGLRKP 504
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 446 FV-QAIAQCKHTPPAPKEDDPLKKTTFEVLNKEKFSADLDQPLGRKFLVQAGLRKP 504
QY 480 --TIGPKRSAPSATSSKPAKR 500
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 480 --TIGPKRSAPSATSSKPAKR 500
QY 505 KFTLGKRR-ATPTSSSTSTYAKR 526
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 505 KFTLGKRR-ATPTSSSTSTYAKR 526

```

```

RESULT 9
US-10-301-260A-1
; Sequence 1, Application US/10301260A
; Publication No. US20030118609A1
; GENERAL INFORMATION:
; APPLICANT: Harrison, Stephen
; APPLICANT: Chen, Xiaojiang
; TITLE OF INVENTION: Vaccine compositions and methods for human papilloma
; TITLE OF INVENTION: virus
; FILE REFERENCE: Harvard/Harrison 12687/1123
; CURRENT APPLICATION NUMBER: US/10/301,260A
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: 09/520,822
; PRIOR FILING DATE: 2000-03-18
; PRIOR APPLICATION NUMBER: 60/125208
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/148544
; PRIOR FILING DATE: 1999-08-12
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO: 1
; LENGTH: 531
; TYPE: PRT
; ORGANISM: Human papillomavirus type 16
US-10-301-260A-1

```

```

Query Match 65.9%; Score 1801; DB 15; Length 531;
Best Local Similarity 66.0%; Pred. No. 2,8e-177;
Matches 332; Conservative 69; Mismatches 96; Indels 6; Gaps 5;
QY 1 MALMRSDMTVYLPPEPSAVRVNTDDYTRTSIFYHAGSSRLTGVNPFYRVVAGGANKQ 60
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 27 MSIMLPSEAVTVLPPEPSAVRVNTDDYTRTSIFYHAGSSRLTGVNPFYRVVAGGANKQ 60
QY 61 DIPKVSAYQYRFRVQLPDPNKGFLPDNSIYNFETQRLVWACAGVEIGRGQPLGVLGSH 120
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 DIPKVSAYQYRFRVQLPDPNKGFLPDNSIYNFETQRLVWACAGVEIGRGQPLGVLGSH 120
QY 87 LVPKVSGLQYRFRVRLPDPNKGFLPDNSIYNFETQRLVWACAGVEIGRGQPLGVLGSH 146
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 87 LVPKVSGLQYRFRVRLPDPNKGFLPDNSIYNFETQRLVWACAGVEIGRGQPLGVLGSH 146
QY 121 PFYNTKLDDESSHAATSNVSEYRDVNVSDYKOTQCLIGCAPAIGEHNAKGTACKSRPL 180
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 PFYNTKLDDESSHAATSNVSEYRDVNVSDYKOTQCLIGCAPAIGEHNAKGTACKSRPL 180
QY 147 PLINKLDDTENASAVYAMAGVDRNRECI SMYKOTQCLIGCKPPIEBHMGKSPCTNVAV 206
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 147 PLINKLDDTENASAVYAMAGVDRNRECI SMYKOTQCLIGCKPPIEBHMGKSPCTNVAV 206
QY 181 SOGDCEPPELKNVTLEDGMDVDTGYGAMPESTLQDKCEVPLDICSICKPYDYLQMSAD 240
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 SOGDCEPPELKNVTLEDGMDVDTGYGAMPESTLQDKCEVPLDICSICKPYDYLQMSAD 240
QY 207 NPEDCEPPELKNVTLEDGMDVDTGYGAMPESTLQDKCEVPLDICSICKPYDYLQMSAD 266
    :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 207 NPEDCEPPELKNVTLEDGMDVDTGYGAMPESTLQDKCEVPLDICSICKPYDYLQMSAD 266

```



LENGTH: 508 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 1  
US-09-822-662-1

Query Match 48.5%; Score 1324.5; DB 10; Length 508;  
Best Local Similarity 49.2%; Pred. No. 5,8e-128;  
Matches 249; Conservative 88; Mismatches 158; Indels 11; Gaps 8;

QY 1 MALMRSDNTVYLPPPSVARVNTDDYVTRTSIFYHAGSSRLLTGPNPFYRPAGGANKQ 60  
DB 7 MAVMLPAQNKFTYLPQPSSTVLTDEYKSTINTIFYHASSRLLTVGHPEI-----YKE 55  
QY 61 DIPKVASQYRVFRVQLPDPNKGFLPDNSIYNPEORLWACAGVITIGRQPLGVLSGH 120  
DB 65 TLPKVSNAFRVFRVADPNRPAFGDKALFNPETRLVWGLGIGRQPLGIGITGH 124  
QY 121 PFVKKLDPTES--SHAATSNVSEEDVRDNVSVYKQOTLCIGCAPAIGEMAKGTACKSRP 179  
DB 125 PLINKDDADENPTNTINTHANGDSRONTAFDAKQTMFLVGCIPASGEHNT--SRKCPGEQ 183  
QY 180 LSGDCPPLKNTVLEDDMDVDTGYGAMDFTLDPTKCEVPLDICSICKPYDYLQMSA 239  
DB 184 VKIGDCPVRVYMIESYIEDGDMDMDIGFAMDFAQQDSDVPLDVYQATCKPYDIRMNH 243  
QY 240 DPGDSMEFCLREQGFARHFWNRAGTMGD--TVPSQLYIKGTG--KRASGSCVYSPSPS 296  
DB 244 EAVGNMFEFFARREQWYTHFFTRGGSVDGKEAVPSQSLYLTADAERTLTATNTYVGTPS 303  
QY 297 GSIVTSDQLFNKPYMLHKAQGNHNGICMHNOLFTVVDTRSTNTICASTQSPYVGQY 356  
DB 304 GSNVSSDVLQFNSTYLGQGGNNGIGNRNOLFIVGDNTRSTISI--SMKNNSTY 361  
QY 357 DATAKQYRHVEEYDLOFTFOLCTTLTADVMSYHSNMSSTLEDNMNFGVPPPTTSV 416  
DB 362 SNANFDLRLHEEFLEFIVQLCKVLTPENLAYIHTMDPNILEDNMQLSVSQPPIPLE 421  
QY 417 DRYRFV--GSVAITICQDAAPAEKNDPYDLKFMNVDLKEKFSLDLDQYRGRREYVQAGL 475  
DB 422 DQIRFGSSILAKCPQADPEPOTDPSQKFEWDLTERMSQLDQFPLGRKFLYQSGM 481  
QY 476 -RRKPTIGPRKRSAPSATSSKPAKR 500  
DB 482 TORATSTTKRKTVAVSTSAKRRK 507

## RESULT 12

US-09-134-377-2  
Sequence 2, Application US/09134377  
Patent No. US20020015710A1  
GENERAL INFORMATION:  
APPLICANT: SCHLEGEL, C. Richard  
APPLICANT: JENSON, A. Bennett  
APPLICANT: GHIM, Shin-Je  
TITLE OF INVENTION: PAPILLOMAVIRUS VACCINES  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: BURNS, DOANE, SWECKER & MATHEIS  
STREET: P.O. Box 1404  
CITY: Alexandria  
STATE: Virginia  
COUNTRY: United States  
ZIP: 22313-1404  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/134, 377

FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/724, 281  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Teskin, Robin L.  
REGISTRATION NUMBER: 35,030  
REFERENCE/DOCKET NUMBER: 010091-015  
TELEPHONE: (703) 836-6620  
TELEFAX: (703) 836-2021  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 503 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-134-377-2

Query Match 47.5%; Score 1298; DB 9; Length 503;  
Best Local Similarity 48.2%; Pred. No. 3.2e-125;  
Matches 247; Conservative 92; Mismatches 153; Indels 20; Gaps 8;

QY 1 MALMRSDNTVYLPPPSVARVNTDDYVTRTSIFYHAGSSRLLTGPNPFYRPAGGANKQ 60  
DB 1 MAVMLPAQNKFTYLPQPSSTVLTDEYKSTINTIFYHASSRLLTVGHPEI-----YKE 55  
QY 61 D-----IPKVASQYRVFRVQLPDPNKGFLPDNSIYNPEORLWACAGVEIGRQPLG 114  
DB 56 ERSEEVIVPKNVSPNOYRVEFLLPDPNNFAGDKSLFDPKEKRLVWGLRGLFGRQPLG 115  
QY 115 VGSIGHPYVKLDDE--SSHAATSNVSEEDVRDNVSVYKQOTLCIGCAPAIGEMAKG 172  
DB 116 ISYGHPTFDRIYNDVDPNKNLHAGGGTDSRVNMGIDPKQOTQMFMIQCKPALGEMSLT 175  
QY 173 TACKSRPLSGDCPPLKNTVLEDDMDVDTGYGAMDFTLDPTKCEVPLDICSICKPY 232  
DB 176 RMCTGVNHTAGCQPPLELRNTTIEDGMDVIGFAMDFAKQYHKGVSVPDIYNSAKTY 235  
QY 233 DYLQMSADPYGDSMEFCLREQGFARHFWNRAGTMG--DVPQSLYIKGTGMRASPGSCY 291  
DB 236 DYLMANENEPQCFEVRREQLYARHMSRSGTGLEPVPKDTY--ATREDNNTGTNY 293  
QY 292 SPSPSGSIYSDQLFNKPYMLHKAQGNHNGICMHNOLFTVVDTRSTNTICASTQ--S 350  
DB 294 FSTPSSGLSVSSGOLNRPYWIORSQGNKNGIAMGNOLFTVVDTRGTPLTINIGQDQK 353  
QY 351 PVBQYDATKFKQYSHVEEYDLOFTFOLCTTLTADVMSYHSNMSSTLEDNMNFGVPP 410  
DB 354 PEQGNVYVPSYRFLRYLREVEEYISIVQLCKVLTPENLAIHTMDPNILEDNMHLYVTP 413  
QY 411 PITSVDYTRFVQSVAILTCQDAAPAEKNDPYDLKFMNVDLKEKFSLDLDQYRGRKFL 470  
DB 414 SGT-LDDYTRYINSILTKCTNIPKTNVDPFDFEWDLDKMTQDQPLGRKFL 472  
QY 471 VQAGLRKPTIGPRKRSAPSATSSKPAKRV 502  
DB 473 FQTNVLRKRV--KVNSTSHVYKRAVKKR 502

## RESULT 13

US-09-822-662-2  
Sequence 2, Application US/09822662  
Publication No. US20020197264A1  
GENERAL INFORMATION:  
APPLICANT: SCHLEGEL, C. Richard  
APPLICANT: JENSON, A. Bennett  
APPLICANT: GHIM, Shin-Je  
TITLE OF INVENTION: PAPILLOMAVIRUS VACCINES  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:

```

ADDRESS: BURNS, DOANE, SNECKER & MATHIS
STREET: P.O. Box 1404
City: Alexandria
STATE: Virginia
COUNTRY: United States
ZIP: 22133-1404

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/822,662
FILING DATE: 02-Apr-2001
CLASSIFICATION: <unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/724,281
FILING DATE: <unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Teskin, Robin L.
REGISTRATION NUMBER: 35,030
REFERENCE/DOCKET NUMBER: 010091-015
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-6620
TELEFAX: (703) 836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 503 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-822-662-2

Query Match      47.5%; Score 1298; DB 10; Length 503;
Best Local Similarity 48.2%; Pred. No. 3.2e-125;
Matches 247; Conservative 92; Mismatches 153; Indels 20; Gaps 8;

OY 1 MALMRPSDNVYLLPPPSVARYVNTDDYVTRTSIFYHAGSSRLTLVGNPFYRVPAAGGNKQ 60
DB 1 MAVVLPQNKRYLLPPQSTKVLSTDEVYSRTNIFYHASSERLLTVGHFYEI-----YKE 55
OY 61 D-----IPKSAVOYRVFVOLPDPNKGFLPDSIYVPEYQRLVMACAGVEIGRGQPLG 114
DB 56 ERSESVIYPKVSPQYRFFLLPDPNNEAFGDSLEPKEKRLVWGLGLEIGRGQPLG 115
OY 115 VGLSGHPYKRLDYTE--SSHAATSNVSEVDYVNDVYDKOTQCLIGCAPAIGHMAKG 172
DB 116 ISVIGHPTFDYRNVENENKMLAGHGGETSRVNMGLDPRKOTQFMIGCKRALKGHSGLT 175
OY 173 TACKSRPLSGDCRPLELKNYLEDGDMVDTGYGAMDFSTLDTKCEVPLDICOISICKYR 232
DB 176 RMCGOVHTAGQCPRLELNTTIEDGMVDIGFGAMDFKALQHYKSGVPDIIIVSACKYR 235
OY 233 DYLMSADPYGDSMFCRLRREOLFARHFWNRAGTMG-DIVPOSILYIKTGMRASPGSCVY 291
DB 236 DYLMAANPEYGRCFEVRREOLYARHIMSRSGTGLEPVPKDY--ATREDNNIGTNY 293
OY 292 SPSPSGIVTSDOLFENKPYMLHKAQGNHNCICWHNOLFVTVVDTTRSTNLTICASTQ-S 350
DB 294 FSTPSGSLVSSGQLFNRPIYIORSQGNNGIANGNOLFVTVVDTTRKTPPLTINIGQDK 353
OY 351 PVPQOYATKPKQSRHYEEDLOFIQOLCTITTLADVMSYIHSNNSILEDWNGVYPP 410
DB 354 PEEGNYVSSRYTLRHVEEYEVSLIYQLCKVLSPEMLAIITHMDPIIIDMHNIVP 413
OY 411 PTTSLVDTYRVQSAITCQDAAPAEKNDPYDKLKFNVLDLKEKFSJLDLQYPLGRKFL 470
DB 414 SGT-LDDTYRYINSILCTPINIPKTVNDPFRDFKEWEVDLKDMEQLDQTPLGKFL 472
OY 471 VQAGIRKRPITGPRKRSAPSATSSKPAKRYR 502

```

```

DB 473 FQNVLRRLRSV--KVSTSHVSKRAVKRRR 502

RESULT 14
US-10-305-765-2
Sequence 2, Application US/10305765
Publication No. US20030175907A1
GENERAL INFORMATION:
APPLICANT: FRAZER, Ian Hector
APPLICANT: ZHOU, Jian
TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND METHOD FOR SELECTIVELY
FILE REFERENCE: 210338.0001/1US
CURRENT FILING DATE: 2002-11-27
PRIOR APPLICATION NUMBER: US/10/305,765
PRIOR FILING DATE: 2000-01-07
PRIOR APPLICATION NUMBER: PCT/AU98/00530
PRIOR FILING DATE: 1998-07-09
PRIOR APPLICATION NUMBER: AU P07765
PRIOR FILING DATE: 1997-07-09
PRIOR APPLICATION NUMBER: AU P09467
PRIOR FILING DATE: 1997-09-11
NUMBER OF SEQ ID NOS: 219
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 2
LENGTH: 495
TYPE: PRT
ORGANISM: Bovine papillomavirus type 1
FEATURE:
OTHER INFORMATION: LI open reading frame (wild-type)
US-10-305-765-2

Query Match      46.9%; Score 1283; DB 12; Length 495;
Best Local Similarity 48.3%; Pred. No. 1.1e-123;
Matches 243; Conservative 90; Mismatches 156; Indels 14; Gaps 7;

OY 1 MALMRPSDNVYLLPPPSVARYVNTDDYVTRTSIFYHAGSSRLTLVGNPFYRVPAAGGNKQ 60
DB 1 MALMOQGR-LYLPPTPVSKVLCSEYVQKRSIFYHETERLLTIGHYPVSGA---K 56
OY 61 DIPKSAVOYRVFVOLPDPNKGFLPDSIYVPEYQRLVMACAGVEIGRGQPLGVLGSH 120
DB 57 TVPKVASNQYVVEFIQLLDPNQLPDRYVHNSPKERLWVPVIGVYSGQPLGTVTGH 116
OY 121 PFYKRLDTESSHAATSNVSEVDYVNDVYDKOTQCLIGCAPAIGHMAKGTRCKSRPL 180
DB 117 PTFNALDAENVNKNVYQTDDRKQGLAKQOQIILLGOTPAEGYWTTPARPCVTDRL 176
OY 181 SQGCRPLELKNYLEDGDMVDTGYGAMDFSTLDTKCEVPLDICOISICKYRPTQLMSAD 240
DB 177 ENGACPRLELKNKIEEGDMIEIGFANFEINASKSDPLDIONEICLYPDYLNKAE 236
OY 241 PYGSMFECRLRREOLFARHFWNRAGTMGVPOSILY--KGTGMRASPGSCVYSPSPG 297
DB 237 AAGSMFEPFAKKEVYVYRHITRGSGSEKAPTFYTKNNKGDAITLIPS--VHFGSPG 294
OY 298 SIYVSDOLFENKPYMLHKAQGNHNCICWHNOLFVTVVDTTRSTNLTICASTQSPVPCOYD 357
DB 295 SLVSTDNQIENRPYMLRQAGMNGNIMNLLFLTVGDNTRGTLTISVSDGTPLEIYD 354
OY 358 ATKRKQYSRHYEEDLOFIQOLCTITTLADVMSYIHSNNSILEDWNGVYPPPTSLYD 417
DB 355 SSKFNHYRHHEEYKLIFFIELCSEVETAGVSHLQGLMSVLENWEIGVQPPSSILED 414
OY 418 TYRVQSAITCQDAAPAEKNDPYDKLKFNVLDLKEKFSJLDLQYPLGRFLVQAGLR 477
DB 415 TYRIESPATKCAENVLPK-EDPYAGKFWNIDLKELSLDLDQFLGRFLVQAGAGC 473
OY 478 KPTIGPRKRSAPSATSSKPAK 500
DB 474 STV---KKRRI-SQKTSKPAK 492

```

## RESULT 15

US-10-305-765-4

; Sequence 4, Application US/10305765

; Publication No. US20030175907A1

; GENERAL INFORMATION:

; APPLICANT: FRAZER, Ian Hector

; APPLICANT: ZHOU, Jian

; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND METHOD FOR SELECTIVELY

; TITLE OF INVENTION: EXPRESSING A PROTEIN IN A TARGET CELL OR TISSUE

; FILE REFERENCE: 210338.0001/105

; CURRENT APPLICATION NUMBER: US/10/305,765

; CURRENT FILING DATE: 2002-11-27

; PRIOR APPLICATION NUMBER: US/09/479, 645

; PRIOR FILING DATE: 2000-01-07

; PRIOR APPLICATION NUMBER: PCT/AU98/00530

; PRIOR FILING DATE: 1998-07-09

; PRIOR APPLICATION NUMBER: AU P07765

; PRIOR FILING DATE: 1997-07-09

; PRIOR APPLICATION NUMBER: AU P09467

; PRIOR FILING DATE: 1997-09-11

; NUMBER OF SEQ ID NOS: 219

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 4

; LENGTH: 495

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: Description of Artificial Sequence: Bovine

; OTHER INFORMATION: papillomavirus type 1 L1 open reading frame

; FEATURE:

; OTHER INFORMATION: wild-type codons replaced with synonymous codons

; OTHER INFORMATION: used at relatively high frequency by human genes

US-10-305-765-4

Query Match

Best Local Similarity 46.9%; Score 1283; DB 12; Length 495;

Matches 243; Conservative 90; Mismatches 156; Indels 14; Gaps 7;

```
QY 1 MALMPSPDWTYYLPPPSVARYVNTDYYRTSIFYHAGSSRLTLVGNPYFRYPAGGCKNO 60
DB 1 MALMOGOK-LYLPPTPSKYVLCSEYTVQRKSIFFHAETERLITIGHPTVPVSIQA--K 56
QY 61 DIPKVSAYQYRVFVQLPDPNKEGLPDSINYPETQRLVACAGVEIGRQPLGYGLSGH 120
DB 57 TVPKVSANQYRVFKIQLPDPOFALPDRTVHNPSEKRLVPIGVVSRGQPLGTVTGH 116
QY 121 PPKNLDTESHAATSNVSEVDYRNVSYDYKOTLCITGAPALGEHMAKGTACKSRPL 180
DB 117 PTFNALDLAENVNRKVTQTTDRKOTGIDAKOOOILLGCTPAEGEYWTARPCVTDRL 176
QY 181 SOGDPPPELKNVTLEDGMDVDTGYGAMDFSLDPTKCEVPLDIOGOSTICKYPPDYLOMSAD 240
DB 177 ENGACPPPELKNKHTLEDGMDMEIGFANFEKINASKSDPLDIONEICLYPDYLKMAED 236
QY 241 PYGDMFPCFLRREQLFARHFNNRAGTMDPTVPQSLYI--KGTGMRASPSGCVSPSPSG 297
DB 237 AAGNMGFFARKEQYVVRIRIWRGSEKAPPTDLYLKNKGDATLKIPS--VHGSPPSG 294
QY 298 SLVTSDSQLFNKPYVLHKAQGHNGICWHNQLFTVVDTRTSNLTICASTQSPVPGQYD 357
DB 295 SLVSTDNQIFNRPYVLFRAQGMNNGIAMNLLFLTVGDNTRGTNLTISVADGTPLTEYD 354
QY 358 ATKFKOYSRHVEYDLOFIFOLCTITLTDVMSYIHSNMSSTLEDMNFQVPPPTSLVD 417
DB 355 SKKENYVHRHMEYKLAFTLECSVETIAQTVSHLQGLAPSVLENMEIGVQPTTSILED 414
QY 418 TTRFYQVAITQCKDAAPAKENKDPDKLTFMNVVDLKEKFSLDLQYPLGRKFLVQAGLR 477
DB 415 TVRYTESPATKCAANVTPAK-EDPYAGCFEWNIDLEKELSLDLQDPLGRRLAQAGACC 473
QY 478 KPTIGPKRKSAPSATTSSKPAKR 500
```

DB 474 STV---RKRRI-SQRTSSKPAKR 492

Search completed: October 17, 2003, 11:15:36  
Job time : 49.6622 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2003, 10:57:04 ; Search time 22.5217 Seconds  
(without alignments)  
952,486 Million cell updates/sec

Title: US-08-913-644-2

Perfect score: 2733  
Sequence: 1 MALMRPSDNTVTLPPPSVAR.....APSATTSSKPAKRVRRANK 507

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfile.s1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2733	100.0	507	2	US-08-409-122-2
2	2733	100.0	507	2	US-08-408-669-2
3	2705	99.0	568	4	US-09-520-822A-3
4	1826	66.8	505	3	US-08-920-822A-2
5	1826	66.8	505	4	US-08-944-368A-2
6	1826	66.8	505	4	US-09-820-764-2
7	1801	65.9	531	4	US-09-986-118A-2
8	1789.5	65.5	497	4	US-09-520-822A-1
9	1789.5	65.5	499	4	US-09-413-611A-6
10	1768.5	64.7	532	3	US-08-737-336-6
11	1742.5	63.8	500	2	US-08-987-519-2
12	1726.5	63.2	501	2	US-08-987-519-1
13	1679	61.4	501	2	US-08-987-519-3
14	1324.5	48.5	508	2	US-08-724-281-1
15	1324.5	48.5	508	4	US-09-134-377-1
16	1298	47.5	503	2	US-08-724-281-2
17	1298	47.5	503	4	US-09-134-377-2
18	1283	46.9	495	4	US-09-479-645A-2
19	1283	46.9	495	4	US-09-479-645A-4
20	1271	46.5	964	4	US-08-484-791-2
21	604	22.1	132	4	US-09-000-266-29
22	604	22.1	132	4	US-09-628-099-29
23	604	22.1	132	4	US-10-056-360-29
24	604	22.1	132	4	US-10-056-359-29
25	510	18.7	128	4	US-09-319-056B-14
26	493	18.0	128	4	US-09-319-056B-17
27	487.5	17.8	127	4	US-09-319-056B-11

28	423.5	15.5	145	4	US-09-485-973-10	Sequence 10, Appl
29	396	14.5	136	4	US-09-319-056B-8	Sequence 8, Appl
30	382.5	14.0	131	4	US-09-319-056B-23	Sequence 23, Appl
31	379.5	13.9	145	4	US-09-485-973-8	Sequence 8, Appl
32	374	13.7	144	4	US-09-000-266-5	Sequence 5, Appl
33	374	13.7	144	4	US-09-628-099-5	Sequence 5, Appl
34	374	13.7	144	4	US-10-056-360-5	Sequence 5, Appl
35	374	13.7	144	4	US-10-056-359-5	Sequence 5, Appl
36	367	13.4	128	4	US-09-319-056B-5	Sequence 5, Appl
37	363	13.3	136	4	US-09-000-266-8	Sequence 8, Appl
38	363	13.3	136	4	US-09-628-099-8	Sequence 8, Appl
39	363	13.3	136	4	US-10-056-360-8	Sequence 8, Appl
40	363	13.3	136	4	US-10-056-359-8	Sequence 8, Appl
41	363	13.3	138	4	US-09-319-056B-2	Sequence 2, Appl
42	359.5	13.2	139	4	US-09-000-266-2	Sequence 2, Appl
43	359.5	13.2	139	4	US-09-628-099-2	Sequence 2, Appl
44	359.5	13.2	139	4	US-10-056-360-2	Sequence 2, Appl
45	359.5	13.2	139	4	US-10-056-359-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-08-409-122-2, Application US/08409122

Patent No. 5820870

GENERAL INFORMATION:

APPLICANT: JOYCE, JAMES G.

APPLICANT: GEORGE, HUGH A.

APPLICANT: HOFMANN, KATHRYN J.

APPLICANT: JANSSEN, KATHRYN U.

APPLICANT: NEESPER, MICHAEL P.

TITLE OF INVENTION: DNA ENCODING HUMAN PAPILLOMAVIRUS TYPE 18 VACCINE

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: CHRISTINE E. CARTY - MERCK & CO., INC.

STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000

CITY: RAHWAY

STATE: NJ

COUNTRY: US

ZIP: 07065-0907

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 1.5

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/409,122

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/408,669

FILING DATE: 22-MAR-1995

ATTORNEY/AGENT INFORMATION:

NAME: CARTY, CHRISTINE E

REGISTRATION NUMBER: 36,099

REFERENCE/DOCKET NUMBER: 19425

TELECOMMUNICATION INFORMATION:

TELEPHONE: 908-594-6734

TELEFAX: 908-594-4720

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 507 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

HYPOTHETICAL: NO

ANTI-SENSE: NO

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

US-08-409-122-2

Query Match 100.0%; Score 2733; DB 2; Length 507;  
 Best Local Similarity 100.0%; Pred. No. 2, 2e-278;  
 Matches 507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALMRPSDNTVYLLPPPSVAVRVNTDVTYRTSIFYHAGSSRLTLVGNPFRVPAAGGNKQ 60  
 |||||||  
 DB 1 MALMRPSDNTVYLLPPPSVAVRVNTDVTYRTSIFYHAGSSRLTLVGNPFRVPAAGGNKQ 60  
 |||||||  
 QY 61 DIPKVSAYQYRVFVQVLPDPNKKFGLPDNSIYNPETORLVNACAGVEIGRQPLGVLSGH 120  
 |||||||  
 DB 61 DIPKVSAYQYRVFVQVLPDPNKKFGLPDNSIYNPETORLVNACAGVEIGRQPLGVLSGH 120  
 |||||||  
 QY 121 PFYKKLDPTSSHAATSNVEDVDVNSVDYKQTOCLILGCAPAIGEHMAKGTACKSRPL 180  
 |||||||  
 DB 121 PFYKKLDPTSSHAATSNVEDVDVNSVDYKQTOCLILGCAPAIGEHMAKGTACKSRPL 180  
 |||||||  
 QY 181 SOGDCPPELEKNTVLEGGDMVDYGYGAMDSTLDTCEVPLDTCOSICKYPPYLQMSAD 240  
 |||||||  
 DB 181 SOGDCPPELEKNTVLEGGDMVDYGYGAMDSTLDTCEVPLDTCOSICKYPPYLQMSAD 240  
 |||||||  
 QY 241 PYGDSMFCLRRBOLFARHFMNRRAGTMDVYPOSILYKGGMRASPGSCYSPSPSGSIY 300  
 |||||||  
 DB 241 PYGDSMFCLRRBOLFARHFMNRRAGTMDVYPOSILYKGGMRASPGSCYSPSPSGSIY 300  
 |||||||  
 QY 301 TSDSOLFENKPYWLKKAAGHNNICWNNQLFVTVVDTTRSTNLITICASTQSPVPGQYDATK 360  
 |||||||  
 DB 301 TSDSOLFENKPYWLKKAAGHNNICWNNQLFVTVVDTTRSTNLITICASTQSPVPGQYDATK 360  
 |||||||  
 QY 361 FKQSRHVEEYDLOFTIOLCTITLTADVMSYIHSNSSLIEDMNFVGPPTTSLVDTYR 420  
 |||||||  
 DB 361 FKQSRHVEEYDLOFTIOLCTITLTADVMSYIHSNSSLIEDMNFVGPPTTSLVDTYR 420  
 |||||||  
 QY 421 FVQSVATTCQKDAAPAKNKPYDKLKFWNVDLKEKFSILDQYPLGKRFVQAGLRKRP 480  
 |||||||  
 DB 421 FVQSVATTCQKDAAPAKNKPYDKLKFWNVDLKEKFSILDQYPLGKRFVQAGLRKRP 480  
 |||||||  
 QY 481 IGPKRSAPSATYSSKPAKRVVRARK 507  
 |||||||  
 DB 481 IGPKRSAPSATYSSKPAKRVVRARK 507  
 |||||||

RESULT 2

US-08-408-669-2

; Sequence 2, Application US/08408669  
 ; Patent No. 5840306

; GENERAL INFORMATION:

; APPLICANT: JOHANN, KATHRYN J.

; APPLICANT: JENSEN, MICHAEL P.

; TITLE OF INVENTION: DNA ENCODING HUMAN PAPILLOMAVIRUS TYPE 18

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000

; CITY: RAHWAY

; STATE: NJ

; COUNTRY: US

; ZIP: 07065-0907

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq Version 1.5

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/408,669

; FILING DATE: 22-MAR-1995

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

NAME: CARTY, CHRISTINE E  
 REGISTRATION NUMBER: 36,099  
 REFERENCE/DOCKET NUMBER: 19424  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 908-594-6734  
 TELEFAX: 908-594-4720

; TELEX:

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 507 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; FRAGMENT TYPE: N-terminal

; ORIGINAL SOURCE:

US-08-408-669-2

Query Match 100.0%; Score 2733; DB 2; Length 507;  
 Best Local Similarity 100.0%; Pred. No. 2, 2e-278;  
 Matches 507; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MALMRPSDNTVYLLPPPSVAVRVNTDVTYRTSIFYHAGSSRLTLVGNPFRVPAAGGNKQ 60  
 |||||||  
 DB 1 MALMRPSDNTVYLLPPPSVAVRVNTDVTYRTSIFYHAGSSRLTLVGNPFRVPAAGGNKQ 60  
 |||||||  
 QY 61 DIPKVSAYQYRVFVQVLPDPNKKFGLPDNSIYNPETORLVNACAGVEIGRQPLGVLSGH 120  
 |||||||  
 DB 61 DIPKVSAYQYRVFVQVLPDPNKKFGLPDNSIYNPETORLVNACAGVEIGRQPLGVLSGH 120  
 |||||||  
 QY 121 PFYKKLDPTSSHAATSNVEDVDVNSVDYKQTOCLILGCAPAIGEHMAKGTACKSRPL 180  
 |||||||  
 DB 121 PFYKKLDPTSSHAATSNVEDVDVNSVDYKQTOCLILGCAPAIGEHMAKGTACKSRPL 180  
 |||||||  
 QY 181 SOGDCPPELEKNTVLEGGDMVDYGYGAMDSTLDTCEVPLDTCOSICKYPPYLQMSAD 240  
 |||||||  
 DB 181 SOGDCPPELEKNTVLEGGDMVDYGYGAMDSTLDTCEVPLDTCOSICKYPPYLQMSAD 240  
 |||||||  
 QY 241 PYGDSMFCLRRBOLFARHFMNRRAGTMDVYPOSILYKGGMRASPGSCYSPSPSGSIY 300  
 |||||||  
 DB 241 PYGDSMFCLRRBOLFARHFMNRRAGTMDVYPOSILYKGGMRASPGSCYSPSPSGSIY 300  
 |||||||  
 QY 301 TSDSOLFENKPYWLKKAAGHNNICWNNQLFVTVVDTTRSTNLITICASTQSPVPGQYDATK 360  
 |||||||  
 DB 301 TSDSOLFENKPYWLKKAAGHNNICWNNQLFVTVVDTTRSTNLITICASTQSPVPGQYDATK 360  
 |||||||  
 QY 361 FKQSRHVEEYDLOFTIOLCTITLTADVMSYIHSNSSLIEDMNFVGPPTTSLVDTYR 420  
 |||||||  
 DB 361 FKQSRHVEEYDLOFTIOLCTITLTADVMSYIHSNSSLIEDMNFVGPPTTSLVDTYR 420  
 |||||||  
 QY 421 FVQSVATTCQKDAAPAKNKPYDKLKFWNVDLKEKFSILDQYPLGKRFVQAGLRKRP 480  
 |||||||  
 DB 421 FVQSVATTCQKDAAPAKNKPYDKLKFWNVDLKEKFSILDQYPLGKRFVQAGLRKRP 480  
 |||||||  
 QY 481 IGPKRSAPSATYSSKPAKRVVRARK 507  
 |||||||  
 DB 481 IGPKRSAPSATYSSKPAKRVVRARK 507  
 |||||||

RESULT 3

US-09-520-822A-3

; Sequence 3, Application US/09520822A

; Patent No. 6551597

; GENERAL INFORMATION:

; APPLICANT: Harrison, Stephen

; APPLICANT: Chen, Xiaojiang

; TITLE OF INVENTION: Vaccine compositions and methods for human papilloma

; TITLE OF INVENTION: virus

; FILE REFERENCE: Harvard/Harrison 12687/1120

; CURRENT APPLICATION NUMBER: US/09/520,822A

; CURRENT FILING DATE: 2000-03-08



; PRIOR APPLICATION NUMBER: 60/125208  
 ; PRIOR FILING DATE: 1999-03-18  
 ; PRIOR APPLICATION NUMBER: 60/148544  
 ; PRIOR FILING DATE: 1999-08-12  
 ; NUMBER OF SEQ ID NOS: 4  
 ; SOFTWARE: Patentln Ver. 2.1  
 ; SEQ ID NO 3  
 ; LENGTH: 568  
 ; TYPE: PR1  
 ; ORGANISM: Human papillomavirus type 18  
 US-09-520-822A-3

Query Match 99.0%; Score 2705; DB 4; Length 568;  
 Best Local Similarity 99.0%; Pred. No. 2,3e-275;  
 Matches 502; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 MALMRSDNTVYLPPEPSAVRVNTDDYVTRTSIFYHAGSSRLTLVGNPYRVVAGGNKQ 60  
 DB 62 MALMRSDNTVYLPPEPSAVRVNTDDYVTRTSIFYHAGSSRLTLVGNPYRVVAGGNKQ 121  
 QY 61 DIPKVSAYOYRVFVOLLPPNKEGLPDNSIYNPETORLVACAGVEIGRGQPLGVLSGH 120  
 DB 122 DIPKVSAYOYRVFVOLLPPNKEGLPDTSIYNPETORLVACAGVEIGRGQPLGVLSGH 181  
 QY 121 PFYNKLDITRESSHAATSNVSEVDYRVNVDYKOTQCLIGCAPAIGEHMAKGTACKSRPL 180  
 DB 182 PFYNKLDITRESSHAATSNVSEVDYRVNVDYKOTQCLIGCAPAIGEHMAKGTACKSRPL 241  
 QY 181 SOGDCEPLELKNVLELDGMDVDTGYGAMDSTLQDKCEVPLDICSICKYPPYLQMSAD 240  
 DB 242 SOGDCEPLELKNVLELDGMDVDTGYGAMDSTLQDKCEVPLDICSICKYPPYLQMSAD 301  
 QY 241 PYGDSMEFCLRRQLFARHFNWAGTMDGTVPOSILYIKGTGMASPGSCYSPSGSIY 300  
 DB 302 PYGDSMEFCLRRQLFARHFNWAGTMDGTVPOSILYIKGTGMASPGSCYSPSGSIY 361  
 QY 301 TTSOQLFNKPYWHLKQGHNGICWGNOLFVTVVDTTRSTNLITCSTOSPVVGYDARK 360  
 DB 362 TTSOQLFNKPYWHLKQGHNGICWGNOLFVTVVDTTRSTNLITCSTOSPVVGYDARK 421  
 QY 361 FKOYSRHEVEYDLOFIFOLCTITLTADVMSYIHSNMSILEDNFGVPPPTSLVDYR 420  
 DB 422 FKOYSRHEVEYDLOFIFOLCTITLTADVMSYIHSNMSILEDNFGVPPPTSLVDYR 481  
 QY 421 FVOSVAITCQKDAAPAKNDPYDKLKFVNVDLKEKESLDLDQYPLGRKFLVQAGLRKRP 480  
 DB 482 FVOSVAITCQKDAAPAKNDPYDKLKFVNVDLKEKESLDLDQYPLGRKFLVQAGLRKRP 541  
 QY 481 IGPKRKSAPSATSSKPAKRVVRARAK 507  
 DB 542 IGPKRKSAPSATSSKPAKRVVRARAK 568

RESULT 4  
 US-08-944-368A-2  
 ; Sequence 2, Application US/08944368A  
 ; Patent No. 6228368  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Glasman, et al.  
 ; TITLE OF INVENTION: Papilloma Virus Capsomere Vaccine  
 ; TITLE OF INVENTION: Formulations and Methods of Use  
 ; NUMBER OF SEQUENCES: 28  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray &  
 ; STREET: 233 South Wacker Drive, 6300 Sears Tower  
 ; CITY: Chicago  
 ; STATE: Illinois  
 ; COUNTRY: United States of America  
 ; ZIP: 60606-6402  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentln Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/944,368A  
 ; FILING DATE:  
 ; CLASSIFICATION: 424  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Williams Jr., Joseph A.  
 ; REGISTRATION NUMBER: 38,659  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 312-474-6300  
 ; TELEFAX: 312-474-0448  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 505 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 US-08-944-368A-2

Query Match 66.8%; Score 1826; DB 3; Length 505;  
 Best Local Similarity 66.3%; Pred. No. 5e-183;  
 Matches 333; Conservative 68; Mismatches 97; Indels 4; Gaps 3;

QY 1 MALMRSDNTVYLPPEPSAVRVNTDDYVTRTSIFYHAGSSRLTLVGNPYRVVAGGNKQ 60  
 DB 1 MSIMLPSSEATVYLPPEPSAVRVNTDDYVTRTSIFYHAGSSRLTLVGNPYRVVAGGNKQ 60  
 QY 61 DIPKVSAYOYRVFVOLLPPNKEGLPDNSIYNPETORLVACAGVEIGRGQPLGVLSGH 120  
 DB 61 LVPKVSGLQYRVFRIHLPDNKEGFPDTSFYNDPQRLVACAGVEIGRGQPLGVLSGH 120  
 QY 121 PFYNKLDITRESSHAATSNVSEVDYRVNVDYKOTQCLIGCAPAIGEHMAKGTACKSRPL 180  
 DB 121 PLNLKLDITENASAVANACVNRRECIISMDYKOTQCLIGCKPIGEHMGKSPCTNVAV 180  
 QY 181 SOGDCEPLELKNVLELDGMDVDTGYGAMDSTLQDKCEVPLDICSICKYPPYLQMSAD 240  
 DB 181 NPDCPEPLELKNVLELDGMDVDTGYGAMDSTLQDKCEVPLDICSICKYPPYLQMSAD 240  
 QY 241 PYGDSMEFCLRRQLFARHFNWAGTMDGTVPOSILYIKGTGMASPGSCYSPSGSIY 300  
 DB 241 PYGDSMEFCLRRQLFARHFNWAGTMDGTVPOSILYIKGTGMASPGSCYSPSGSIY 300  
 QY 301 TTSOQLFNKPYWHLKQGHNGICWGNOLFVTVVDTTRSTNLITCSTOSPVVGYDARK 360  
 DB 301 TSDAQLFNKPYWHLKQGHNGICWGNOLFVTVVDTTRSTNLITCSTOSPVVGYDARK 359  
 QY 361 FKOYSRHEVEYDLOFIFOLCTITLTADVMSYIHSNMSILEDNFGVPPPTSLVDYR 420  
 DB 360 FKEYLRHGEYDLOFIFOLCTITLTADVMSYIHSNMSILEDNFGVPPPTSLVDYR 419  
 QY 421 FVOSVAITCQKDAAPAKNDPYDKLKFVNVDLKEKESLDLDQYPLGRKFLVQAGLRKRP 479  
 DB 420 FVTSOAIACQKHPTPKEDPLEKKYFEWNLKEKESADLDQYPLGRKFLVQAGLRKRP 479  
 QY 480 -TIGPKRKSAPSATSSKPAKRVVRARAK 500  
 DB 480 FTLGKRR-ATPTTSTSTAKR 500

RESULT 5  
 US-09-820-764-2  
 ; Sequence 2, Application US/09820764  
 ; Patent No. 635266  
 ; GENERAL INFORMATION:  
 ; APPLICANT: BURGER, Alexander  
 ; TITLE OF INVENTION: PAPILLOMA VIRUS CAPSOMERE VACCINE  
 ; TITLE OF INVENTION: FORMULATIONS AND METHODS OF USE  
 ; NUMBER OF SEQUENCES: 28  
 ; CORRESPONDENCE ADDRESS:

```

ADDRESS: FOLEY & LARDNER
STREET: 3000 K Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.A.
ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/820,764
FILING DATE: 30-Mar-2001
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 09/026,896
FILING DATE: 20-FEB-1998

ATTORNEY/AGENT INFORMATION:
NAME: Sandercock, Colin G.
REGISTRATION NUMBER: 31,298
REFERENCE/DOCKET NUMBER: 37067/102
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-820-764-2

Query Match      66.8%; Score 1826; DB 4; Length 505;
Best Local Similarity 66.3%; Pred. No. 5e-183;
Matches 333; Conservative 66; Mismatches 97; Indels 4; Gaps 3;

QY 1 MALMRPSDNTVYLPSPSVARVYNTDVTSTISFYHAGSSRLTVGNPFRVPAAGGNKQ 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   1 MSMLPSEATVYLPVPVSKVSTDEVYARTNITYHAGTSRLAVGHPYPIKPNNNKI 60

DB 61 DIPVSAIYQYRVERVOLDPNKGRLPDNSIYNPETOQLVWACAGVEIGROPLGVGSGH 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   61 LVPKVSGLQYRVEFIHLDPNKGFGFPDTSFNPTQRLVWACVGEVGRQPLGVGSGH 120

QY 121 PFYVKLDDTSSHAATSNVEDVDVNDVYKQFQCLICAPALIGHMAKGTACKSRPL 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   121 PLNKLDDTERASAYANAGVDNRCISMDYKQFQCLICGKPPIGEHMGKGSFCTNVAV 180

DB 181 SGGDCPPELEKNTVLEDGDVNDGYGAMDFSTLQDTRCEVPLDICSICKYPDYLQMSAD 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   181 NPGCCPPELELINTVYIOGDVNDGFGAMDFTTLOANKSEVPLDICTSICKYPIKKNVSE 240

QY 241 PYGDSMFCRLRREGLEFARHFWNRAGTMDVYPOSLYIKGTGMASPSGCYVSPSGSIY 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   241 PYGDSLFYLRREGDMFVRHLFNRAGAVENPDLYIKSGSTANLASSNYFPSPGSMV 300

DB 301 TSDSQLNKPYYMLKAGGHNNGICMHNQLEFVTVVDTTRSTNLITICASTOSPVPQOYDATK 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   301 TSDQIQRNKPYYMLORAGGHNNGICMGNQLEFVTVVDTTRSTNMSICAAI-SISEETTYKNTN 359

DB 361 FKQYSRHVEEYDLOFIFOLCTITLTVADMSYIHSMSNLSLEDNMFVGPPEPTSLVDTYR 420
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   360 FKEYLRIGEEYDLOFIFOLCTITLTVADMYIHSMSNLSLEDNMFGLQPPPGGTLEDTYR 419

QY 421 FVQSVALTQCKDAAPAKNDPDKLKFWMNDLKKFSLDLDQYPLGKRFYLAQALRRKP- 479
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   420 FVYSQAALQCKHNPAPAKEDPLKRYTFWEVNLKEKFSADLDQFPLGKRFYLAQALRRKP 479

DB 480 -TIGRRKRSAPSATSSKPAKR 500
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   480 FTLLGKRK-ATPTTTSISTTAKR 500

```

```

RESULT 6
US-09-986-118A-2
: Sequence 2, Application US/09986118A
: Patent No. 6562351
: GENERAL INFORMATION:
: APPLICANT: BURGER, Alexander
: TITLE OF INVENTION: PAPILLOMA VIRUS CAPSOMERE VACCINE
: FORMULATIONS AND METHODS OF USE
: NUMBER OF SEQUENCES: 28
: CORRESPONDENCE ADDRESSES:
: ADDRESS: FOLEY & LARDNER
: STREET: 3000 K Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: U.S.A.
: ZIP: 20007-5109

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/986,118A
FILING DATE: 07-No. 6562351-2001
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US 09/026,896
FILING DATE: <Unknown>

ATTORNEY/AGENT INFORMATION:
NAME: Sandercock, Colin G.
REGISTRATION NUMBER: 31,298
REFERENCE/DOCKET NUMBER: 37067/102
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 672-5300
TELEFAX: (202) 672-5399

INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 505 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-986-118A-2

Query Match      66.8%; Score 1826; DB 4; Length 505;
Best Local Similarity 66.3%; Pred. No. 5e-183;
Matches 333; Conservative 66; Mismatches 97; Indels 4; Gaps 3;

QY 1 MALMRPSDNTVYLPSPSVARVYNTDVTSTISFYHAGSSRLTVGNPFRVPAAGGNKQ 60
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   1 MSMLPSEATVYLPVPVSKVSTDEVYARTNITYHAGTSRLAVGHPYPIKPNNNKI 60

DB 61 DIPVSAIYQYRVERVOLDPNKGRLPDNSIYNPETOQLVWACAGVEIGROPLGVGSGH 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   61 LVPKVSGLQYRVEFIHLDPNKGFGFPDTSFNPTQRLVWACVGEVGRQPLGVGSGH 120

QY 121 PFYVKLDDTSSHAATSNVEDVDVNDVYKQFQCLICAPALIGHMAKGTACKSRPL 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   121 PLNKLDDTERASAYANAGVDNRCISMDYKQFQCLICGKPPIGEHMGKGSFCTNVAV 180

DB 181 SGGDCPPELEKNTVLEDGDVNDGYGAMDFSTLQDTRCEVPLDICSICKYPDYLQMSAD 240
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   181 NPGCCPPELELINTVYIOGDVNDGFGAMDFTTLOANKSEVPLDICTSICKYPIKKNVSE 240

QY 241 PYGDSMFCRLRREGLEFARHFWNRAGTMDVYPOSLYIKGTGMASPSGCYVSPSGSIY 300
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   241 PYGDSLFYLRREGDMFVRHLFNRAGAVENPDLYIKSGSTANLASSNYFPSPGSMV 300

DB 301 TSDSQLNKPYYMLKAGGHNNGICMHNQLEFVTVVDTTRSTNLITICASTOSPVPQOYDATK 360
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
   301 TSDQIQRNKPYYMLKAGGHNNGICMHNQLEFVTVVDTTRSTNLTICASTOSPVPQOYDATK 360

```

[illegible]

DB	446	FW-TGALACQCHNTPPAPEDEDLKYYTWEVNLKEKESADDDQFLGKFKLLDGLAKRP	504
QY	480	--TIGPKRRASPSATSSKPAKR 500	
DB	505	KETLKKRK-ATPPTSTSTSTAKR 526	
<p>RESULT 8</p> <p>US-09-413-611A-6</p> <p>Sequence 6, Application US/09413611A</p> <p>Patent No. 6380364</p> <p>GENERAL INFORMATION:</p> <p>APPLICANT: Mueller, Martin</p> <p>APPLICANT: Kast, WlJbe</p> <p>APPLICANT: Nleland, John</p> <p>APPLICANT: Velders, Markwin</p> <p>TITLE OF INVENTION: Chimeric Biotin-Binding Papillomavirus Protein</p> <p>FILE REFERENCE: 202325</p> <p>CURRENT APPLICATION NUMBER: US/09/413.611A</p> <p>CURRENT FILING DATE: 1999-10-06</p> <p>PRIOR APPLICATION NUMBER: US 60/109,510</p> <p>PRIOR FILING DATE: 1998-11-23</p> <p>NUMBER OF SEQ ID NOS: 8</p> <p>SOFTWARE: PatentIn Ver. 2.0</p> <p>SEQ ID NO 6</p> <p>LENGTH: 497</p> <p>TYPE: PRT</p> <p>ORGANISM: Artificial Sequence</p> <p>FEATURE:</p> <p>OTHER INFORMATION: Description of Artificial Sequence: biotin-binding</p> <p>OTHER INFORMATION: Papillomavirus protein 168</p> <p>US-09-413-611A-6</p>			
Query Match	65.5%	Score 1789.5	DB 4; Length 497;
Best Local Similarity	67.1%	Pred. No. 3.4e-179;	
Matches	325;	Conservative 63;	Mismatches 93; Indels 3; Gaps 2.
QY	1	MALMRPSDNTYYLPPPSVARYVNTDDYVTRTSTIFHYAGSSRLTLVGNPFRRYPAGGKQ 60	
DB	1	MSLMLPSATYYLDPVPVSKYSTDEYVARNINYYHAGTSRLNAGHYFFIKRNNKI 60	
QY	61	DIPKYSAYQYRFRQYLDPPNKKFGLPDSIYNPEFQRLVMACAGVEIGRGPGLVGLSGH 120	
DB	61	LVPKSGIQYRFRILHDPNKKFGFPDTSFYNPDQRLVMACAGVEVGRGQPLGVGISGH 120	
QY	121	PFYKKLDTESSHAATSVSESDYRPNVSDVKQPOLCLTGCAPAIAGEMAGTACKSRPL 180	
DB	121	PLNKLKLDTEENASAAANAGVGNRKRCSIMDKYQQLCLIGCKPRPGEHMGKSPCTNAY 180	
QY	181	SGQDCPPLLENTVLEDDDMVDYTGAMDFTLQDTKCEVPLDTCISCKKPYDLOMSAD 240	
DB	181	NPQDCPPLLEINTVYQDDDMVDYTGAMDFTLLQANKSEVPLDTCISCKKPYDIKMYSE 240	
QY	241	PYGDSEMFCLRREQLFARHFNMRAGTMDTVPOSILYIKGTGMRASPGSCVYSPSPSGSIV 300	
DB	241	PYGSLSLFYLLRREQFVHRLFNRRAGAVGENVPDILYIKGSGSTANLASNFFPSPGSMV 300	
QY	301	TSDSOLFNRKPYWLKAQCHNNGICMNLFTVYVDTTSTJMLTICASTQSQSYPGQYDXTK 360	
DB	301	TSDAOLFNRKPYWLRAQCHNNGICMNLFTVYVDTTSTJMSLCAAL-STSETTYKKNTN 359	
QY	361	FKQYSRHEEEDLOFIFQCYTTLTADVASYIHSKNSSLIDDMNGVPPPTSTLVDTYR 420	
DB	360	FKYLIRHGEEDLOFIFQCKITTLTADVMYTIHSNSSLILEDMNGLDPPPGGTLIEDTYR 419	
QY	421	FVQSVATTCQDAAPAEKNDPDKLFKNVNDLKEKESLDDQYPLGRKFLVQAGLRKP- 479	
DB	420	FVTSQALISQCHNTPPAPEDEDLKYYTWEVNLKEKESADDDQFLGKFKLLDGLAKRP 479	
QY	480	-TIG 482	
DB	480	FTLG 483	

```

RESULT 9
US-09-413-611A-8
; Sequence 8, Application US/09413611A
; Patent No. 6380364
; GENERAL INFORMATION:
; APPLICANT: Mueller, Martin
; APPLICANT: Kast, Wiljbe
; APPLICANT: Nieland, John
; APPLICANT: Vellders, Markwin
; TITLE OF INVENTION: Chimeric Biotin-Binding Papillomavirus Protein
; FILE REFERENCE: 202325
; CURRENT APPLICATION NUMBER: US/09/413, 611A
; CURRENT FILING DATE: 1999-10-06
; PRIOR APPLICATION NUMBER: US 60/109,510
; PRIOR FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 8
; LENGTH: 499
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:biotin-binding
; OTHER INFORMATION: papillomavirus protein 169
US-09-413-611A-8

```



```

LOCATION: (0) ... (0)
OTHER INFORMATION: Consensus Sequence
OTHER INFORMATION: At 28 X = K or T
OTHER INFORMATION: At 49 X = Y or F
OTHER INFORMATION: At 53 X = K or R
OTHER INFORMATION: At 54 X = V or A
OTHER INFORMATION: At 119 X = L or F
OTHER INFORMATION: At 131 X = G or S
OTHER INFORMATION: At 132 X = Y or T
OTHER INFORMATION: At 170 X = T or K
OTHER INFORMATION: At 173 X = S or P
OTHER INFORMATION: At 176 X = S or P
OTHER INFORMATION: At 179 X = N or A
OTHER INFORMATION: At 219 X = L or T
OTHER INFORMATION: At 225 X = V or T
OTHER INFORMATION: At 246 X = Y or F
OTHER INFORMATION: At 263 X = T or E
OTHER INFORMATION: At 271 X = D or T
OTHER INFORMATION: At 273 X = L or I
OTHER INFORMATION: At 274 X = V or I
OTHER INFORMATION: At 277 X = G or S
OTHER INFORMATION: At 278 X = N or G
OTHER INFORMATION: At 281 X = S or T
OTHER INFORMATION: At 284 X = A or G
OTHER INFORMATION: At 290 X = H or N
OTHER INFORMATION: At 325 X = H or Q
OTHER INFORMATION: At 346 X = S or T
OTHER INFORMATION: At 347 X = K or T
OTHER INFORMATION: At 348 X = A or T
OTHER INFORMATION: At 366 X = F or Y
OTHER INFORMATION: At 434 X = O or P
OTHER INFORMATION: At 439 X = D or N
OTHER INFORMATION: At 440 X = M or L
OTHER INFORMATION: At 458 X = F or Y
OTHER INFORMATION: At 474 X = T or I
OTHER INFORMATION: At 476 X = A or I
OTHER INFORMATION: At 480 X = I or V
OTHER INFORMATION: At 488 X = P or A
OTHER INFORMATION: At 490 X = T or A
OTHER INFORMATION: At 497 X = T or A
OTHER INFORMATION: At 501 X = K or R
US-08-987-519-3

Query Match      61.4%; Score 1679; DB 2; Length 501;
Best Local Similarity 61.5%; Pred. No. 1.4e-167;
Matches 310; Conservative 69; Mismatches 119; Indels 6; Gaps 4;

QY 3 LMRPSDNTVYLPPPS-VARVNTDDYTRTSTIFYHAGSSRLTVGNPFYRVPAAGGNKOD 61
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 LMRPSDSTVYVPPNPYSKVATDAYKRNIFHASSRLVAGHPYXSTKX--XNKTV 58

QY 62 IPKVSAYQYVFRVYOLPDPNKGFLPDSIYNPETQRLVMACAGVEIGRQPLVGSLGSPH 121
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 59 VPKVSGYQYVFKVYLDPPNKFALPDSLDFPTQRLVMACGLEVRGQPLGVGSGHP 118

QY 122 FYKRLDTESSHATSNVSESDVDNVSVDKQQLCLIGCAPAIGEMWAGTACKSRPLS 181
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 119 XLNKYDVENSGXGKGGPQDNRMVGMQKQQLCVGAPRLEGHWGKGCXQXNTXVQ 178

QY 182 QGDCPPELEKNTVLEDGMDVDTGXAMDFTLQDKCEVPLDQCSICKKPYDLOMSADP 241
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 179 XGDCPPELELTISVYQDDMDVDTGFGAMNFADLQNKSDVFXDIDICGTCKTPIYQMAADP 238

QY 242 YGDSMEFLCRLREOLFARHFNRAAGTMDTVPOSLYIKGTGMRASPGSCVYSPSPSGSI 301
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 239 YGRLFEFLKREQGFARHFNRAAGTMDTVPOSLYIKGTGMRASPGSCVYSPSPSGSIVS 298

QY 302 SDGQLEFKRPYMLKKAQSHNNGICMHNOLFVTVYDTSTNLTICASQSPYPGQYDXTKF 361
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 299 SEQOLEFKRPYMLKKAQSHNNGICMGNLFVTVYDTSTNLTICASQYX--SKTYTNSDY 356

QY 362 KQSRHVEEYDLOFIFOLCTITLADVYSHNSNIILEDQWNGVPPPTTSIVDYRFE 421
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```

```

Db 357 KEIWHAVEEXDLOFIFOLCTITLSAEVMAIHTMNPVLEDMWNGLSPPNGTLEDTRY 416

QY 422 VQSVAITCQKDAAPENKDPYDKLFMNVDLKEKFSIDLQYPLGRKFLVQAGLR-RKPT 480
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 417 VQSVAITCQKPTPEKEXDYPYKXKSEFVNLKEKESSELDQXPLGRFLLQSGVGRGX 476

QY 481 IGPKRSAPSATYSSKPAKRYVR 504
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 477 RTGAKRPVSKXKXAPKRXKTK 500

RESULT 14
US-08-724-281-1
; Sequence 1, Application US/08724281
; Patent No. 5874089
; GENERAL INFORMATION:
; APPLICANT: SCHLEGEL, C. Richard
; APPLICANT: JENSON, A. Bennett
; APPLICANT: GHIM, Shih-Je
; TITLE OF INVENTION: PAPILLOMAVIRUS VACCINES
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P. O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/724,281
; FILING DATE: 01-OCT-1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,691
; FILING DATE: 02-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Teakip, Robin L.
; REGISTRATION NUMBER: 35,030
; REFERENCE/DOCKET NUMBER: 010091-015
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 508 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-724-281-1

Query Match      48.5%; Score 1324.5; DB 2; Length 508;
Best Local Similarity 49.2%; Pred. No. 2.8e-130;
Matches 249; Conservative 88; Mismatches 158; Indels 11; Gaps 8;

QY 1 MALMRPSDNTVYLPPPSVAVVNTDDYTRTSTIFYHAGSSRLTVGNPFYRVPAAGGNKQ 60
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 7 MAWLPAQNKFEYLPQPIITRLSTDEVYTRTNLFYHATSRLLVGHPLPEI--SSNQTV 64

QY 61 DIPKVSAYQYVFRVYOLPDPNKGFLPDSIYNPETQRLVMACAGVEIGRQPLVGSLGSH 120
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 65 TIRKVSNAFRVVRVRRADPNRFAFGDKALFNPETERLVGLGIEIGRQPLGIGTGH 124

QY 121 FYKRLDTESS-SHAATSNVSESDVDNVSVDKQQLCLIGCAPAIGEMWAGTACKSRP 179
    :||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 125 PLNKLDDAENPNYINTYHANGDSRQNTAFDAKQOTQFLVGCIPASEHNT-SRRCGEQ 183

QY 180 LSGDCPPELEKNTVLEDGMDVDTGXAMDFTLQDKCEVPLDQCSICKKPYDLOMSA 239

```

Db	184	VKLGDEPRVQWIESVIEDJEDMDIGFAMDFALQDSDVPLDYQATCKPYDIIRAHN	24.3
Qy	240	DPJGDSMFCLREDOLEFAHFHFNRRAGTGMCD--TVPQSLYITGTC-MRASPOSCYISPSPS	296
Db	244	EAGNSGSMFFEARREDQATYRHFPTRGSGVDKKAAPQSLYLITADAEPRITLATTYVGPBS	303
Qy	297	GSIVTSDSLQFNKPYPYLHKHQAQHNNGNICMHNOLFYTVDTRSNLTICASTOSPVPBGQY	356
Db	304	GSWSSVDVQLFNRSYWLQKQGGQANNIGIRNDLFTVVDNTRGTSLSI--SMKNNASTTY	361
Qy	357	DATKFKQYSRHVEEYDLOEIFQLCTITTLADVMSYIHSMSNLLIEDNFGVPPPTSLY	416
Db	362	SNANFNDPLRHTHEEFQSLFTQLCKVKTLPEMLAVIHTMDPNLIEDMQLSVQPTNPLE	421
Qy	417	DYTRFV-QSVAYITCQKDAAPAEKNKQPYDKLKFANVDLKEKTSLLDQYPLGRKFLVQAGL	475
Db	422	DOYRFLGSSIAKACPEQAPPEPOTBPYQSYKFEWYDLTERMSQDLOFPILGRKFLYQSGM	481
Qy	476	-RRKPTIGPRKSAPGATSSSKPARR 500	
Db	482	TQRTATSSITTKRYAVVSTSAKRKK 507	

```

RESULT 15
US-09-134-377-1
: Sequence 1, Application US/09134377
: Patent No. 6485728
:
GENERAL INFORMATION:
: APPLICANT: SCHLEGEL, C. Richard
: APPLICANT: JENSON, A. Bennett
: APPLICANT: GHIM, Shin-Je
: TITLE OF INVENTION: PAPILLOMAVIRUS VACCINES
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
: STREET: P.O. Box 1404
: CITY: Alexandria
: STATE: Virginia
: COUNTRY: United States
: ZIP: 22313-1404
:
COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/134,377
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/724,281
: FILING DATE:
:
ATTORNEY/AGENT INFORMATION:
: NAME: Teskin, Robin L.
: REGISTRATION NUMBER: 35,030
: REFERENCE/DOCKET NUMBER: 010091-015
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 836-6620
: TELEFAX: (703) 836-2021
:
INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 508 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
US-09-134-377-1

```

Query Match	48.5%;	Score 1324.5;	DB 4;	Length 508;
Best Local Similarity	49.2%;	Pred. No. 2.8e-130;		
Matches 249;	Conservative 88;	Mismatches 158;	Indels 11;	Gaps 8

QY	1	MALMRSDDNVYLPSPSVARVYVTDQYRTSIFFAASSRLTJGNGYFVPAAGGNKQ	60
Db	7	MAWVLPQNKFTYLPQPIRIRLSTDBYVTRTLMFTHAHSERILLVGHFLFEI--SSNOTV	64
QY	61	DIFRVSAYQYRFRVQLPDPNKGFLPDSIYNPEQRLVMAQAGVEIGRGPLVGLSGH	120
Db	65	TIFRVSAPNARVRYRFRVAFDPNRFADKAIENFETERLWVGJRGIEIRGQPLGIGTGH	124
QY	121	PFYNNKLDDPTFS--SHAATSNVSESDVRGNVVDVKOTOLCILGCAPIGGEHMAKGAACKSRP	179
Db	125	PLNNKLDDAENPINTYINTHANGDSKNTLFPDAKQIOMPLGICSTPAGSGHHI--SRRCGEQ	183
QY	180	LSQGDCEPFLKNTVLEDDVNDVTGAGMDFSTLDQKCEYPLDICQSIICKPYLQMSA	235
Db	184	VKIGDCEPRVQMIESTVEDDMDMIGGADFPALQODSDVPLDVQATCKPYIRNNH	243
QY	240	DPIYGDSEFELREREOLEFARHFNWNRAGTMD--TVQSYLYIKGTG--MRAQSPQSYSPSS	296
Db	244	EATGNSKSEFFARREQMTYHFFPRFGGSVDKZKANVQSLXTLTADAEPRTTLATITNYGIPS	300
QY	297	GSIVTSDSOLFNKPWYLHKAQGNHNCICWNOLEFYVVDTRSTNLTICASTQSPVQOY	356
Db	304	GSWSSDYVOLFNRSYWLQGGQGNNGIGWRNOLFTVGGDNTGTSLSI--SMKNAASTTY	364
QY	357	DAIKFKQYSHVHEVQLOLFIQCLCTITTLADVASTIHSNNSIILEDNNFGVPPPTTSLV	411
Db	362	SNANFDELCHTEEPDLSFTYOLCTCKKPLREMLAVIHTHDPDILEEDMOLSVSQPPTMLE	421
QY	417	DIYRFV--QSVAILTCQDAAPAEKNQPYDMLKFWNVNDLKEKESYLDQYPLRKFLYVAGL	475
Db	422	DQYRFGLSSILAACPQAPPEQDTDPYQYKFKWEVDLERHNSQGLDQPLRKFLYQSGM	481
QY	476	-RRKPTIGPKRASPATITSSKPAKR	500
Db	482	TQRAIISSTTKRTVASTAKRRK	507

Search completed: October 17, 2003, 11:03:52  
Job time : 25.5217 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2003, 10:55:49 ; Search time 28.2831 Seconds

(without alignments)  
1723.913 Million cell updates/sec

Title: US-08-913-644-2

Perfect score: 2733  
Sequence: 1 MALMRPSDNTVYLPPPSVAR.....APSATTSSKPARVRVRARK 507

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database: PIR\_76:\*

1: p1r1:\*\n2: p1r2:\*\n3: p1r3:\*\n4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2719	99.5	568	1 P1WL18	L1 protein - human
2	2450	89.6	539	2 S36566	L1 protein - human
3	2155	78.9	505	1 P1WLPR	L1 protein - human
4	2151	78.7	505	1 P1WL39	L1 protein - human
5	1917.5	70.2	532	2 S36554	L1 protein (altern
6	1909	69.8	531	2 S36537	L1 protein - human
7	1879	68.8	508	2 S36508	L1 protein - human
8	1861.5	68.1	503	2 S36549	L1 protein - human
9	1855	67.9	499	2 S36531	L1 protein - human
10	1851	67.7	534	2 S36583	L1 protein - human
11	1834.5	67.1	504	1 P1WL31	L1 protein - human
12	1820.5	66.6	502	1 S36526	L1 protein - human
13	1804.5	66.0	504	1 P1WL51	L1 protein - human
14	1804.5	66.0	524	1 P1WL58	L1 protein - human
15	1801.5	65.9	529	2 S36578	L1 protein - human
16	1801	65.9	531	1 P1WLHS	major capsid prote
17	1797	65.8	505	1 P1WL35	L1 protein - human
18	1791	65.5	499	1 P1WL33	L1 protein - human
19	1789.5	65.5	510	1 S15620	L1 protein - human
20	1784.5	65.3	510	1 S15627	L1 protein - human
21	1782	65.2	502	1 P1WL1C	L1 protein - pygmy
22	1781	65.2	594	2 S36502	L1 protein (altern
23	1771	64.8	505	2 S36589	L1 protein - human
24	1768.5	64.7	500	1 P1WL6	L1 protein - human
25	1766	64.6	505	2 S36520	L1 protein - human
26	1763	64.5	528	2 S36520	L1 protein - human
27	1762	64.5	503	2 S36514	L1 protein - human
28	1759	64.4	499	1 P1WL13	L1 protein - human
29	1757	64.3	502	1 P1WL42	L1 protein - human

30	1733	63.4	501	1 P1WL11	L1 protein - human
31	1724	63.1	502	2 JE0295	L1 protein - human
32	1427	52.2	514	1 P1WL8	L1 protein - human
33	1426.5	52.2	546	2 S36490	L1 protein - human
34	1425.5	52.2	517	2 S36543	L1 protein - human
35	1422.5	52.0	517	2 S36496	L1 protein - human
36	1418	51.9	518	2 S36472	L1 protein - human
37	1416	51.8	507	2 S36595	L1 protein - human
38	1416	51.8	507	2 S36478	L1 protein - human
39	1412.5	51.7	525	1 P1WLBS	L1 protein - human
40	1409	51.6	516	1 P1WL5	L1 protein - human
41	1402	51.3	507	2 S36484	L1 protein - human
42	1389	50.8	514	1 P1WL47	L1 protein - human
43	1379.5	50.5	509	2 S36572	L1 protein - human
44	1378.5	50.4	505	1 P1WLRB	L1 protein - cotto
45	1338.5	49.0	506	1 P1WLBA	L1 protein - bovin

## ALIGNMENTS

## RESULT 1

L1 protein - human papillomavirus type 18

C:Species: human papillomavirus type 18

C:Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 24-Feb-1994

C:Accession: A26251

R: Cole, S.T.; Danos, O.

J. Mol. Biol. 193, 599-608, 1987

A:Title: Nucleotide sequence and comparative analysis of the human papillomavirus typ

A:Reference number: A92937; MID:87283882; PMID:3039146

A:Accession: A26251

A:Molecule type: DNA

A:Residues: 1-568 <COL>

C:Superfamily: papillomavirus L1 protein

C:Keywords: late protein

Query Match 99.5%; Score 2719; DB 1; Length 568;  
Best Local Similarity 99.4%; Pred. No. 5.5e-212;  
Matches 504; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY	1	MALMRPSDNTVYLPPPSVARVNTDYVTRTSIFYHAGSSRLLTVCNPFRRPAGGCKRQ 60	
DB	62	MALMRPSDNTVYLPPPSVARVNTDYVTRTSIFYHAGSSRLLTVCNPFRRPAGGCKRQ 121	
QY	61	DIPKSAVOYRVRVQLPDPNKGFLPDSITVNEGRIVWACGVEIGNGOPRGVLSGH 120	
DB	122	DIPKSAVOYRVRVQLPDPNKGFLPDSITVNEGRIVWACGVEIGNGOPRGVLSGH 181	
QY	121	PFFNKLDDTESSHAATSNVSEEDVDNVSDYKOTOLCIIGCAPAIGEHNAKGTACKSRPL 180	
DB	182	PFFNKLDDTESSHAATSNVSEEDVDNVSDYKOTOLCIIGCAPAIGEHNAKGTACKSRPL 241	
QY	181	SGQDCPEPLEKNTVLEDEGMVDTGYGAMDFSTLQDTKCEVPLDICSICKPYDLOMSAD 240	
DB	242	SGQDCPEPLEKNTVLEDEGMVDTGYGAMDFSTLQDTKCEVPLDICSICKPYDLOMSAD 301	
QY	241	PYDSMFPCRLRRQLFARHFNWAGTMDGPVPSLTIKYKGWAGSPGSCYSPSGSIV 300	
DB	302	PYDSMFPCRLRRQLFARHFNWAGTMDGPVPSLTIKYKGWAGSPGSCYSPSGSIV 361	
QY	301	TSDSQLEFNRPYLHKAQGHNGICWNLFTVVDTRSTNLITICASTOSPVGQYDATK 360	
DB	362	TSDSQLEFNRPYLHKAQGHNGICWNLFTVVDTRSTNLITICASTOSPVGQYDATK 421	
QY	361	EKQYSRHEVEYDQIFOLCTITLTADWASYHSNANSSILEDNNEGVPPPTSLVDYTR 420	
DB	422	EKQYSRHEVEYDQIFOLCTITLTADWASYHSNANSSILEDNNEGVPPPTSLVDYTR 481	
QY	421	FVQSVAITCOCKDAAPENKDPYDKLFMWVNDLKEKSLIDQYPLGRKFLYQAGLRKRP 480	
DB	482	FVQSVAITCOCKDAAPENKDPYDKLFMWVNDLKEKSLIDQYPLGRKFLYQAGLRKRP 541	

QY 481 IGPKRKRSAPATSSSKPAKRVVARK 507  
 Db 542 IGPKRKRSAPATSSSKPAKRVVARK 568

## RESULT 2

S36566

LI protein - human papillomavirus type 45

C:Species: human papillomavirus type 45

C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999

C:Accession: S36566; F44889

R:Deilus, H.; Hofmann, B.

Submitted to the EMBL Data Library, August 1993

A:Description: Primer-directed sequencing of human papillomavirus types.

A:Reference number: S36469

A:Accession: S36566

A:Molecule type: DNA

A:Residues: 1-539 &lt;DEL&gt;

R:van den Brule, A.J.; Snijders, P.J.; Raaphorst, P.M.; Schrijnemakers, H.F.; Deilus, H.

J. Clin. Microbiol. 30, 1716-1721, 1992

A:Title: General primer polymerase chain reaction in combination with sequence analysis

A:Reference number: A44889; MUID:92332706; PMID:1321168

A:Accession: F44889

A:Molecule type: DNA

A:Residues: 366-399 &lt;YAN&gt;

A:Cross-references: GB:S40261; NID:9251692; PIDN:AA22567.1; PID:9251693

A:Experimental source: mucosotropic type 45, cervical smear

A:Note: Sequence extracted from NCBI type 45, cervical smear

C:Superfamily: Papillomavirus LI protein

C:Keywords: late protein

Query Match 89.6%; Score 2450; DB 2; Length 539;

Best Local Similarity 87.3%; Pred. No. 3e-190;

Matches .448; Conservative 30; Mismatches 29; Indels 6; Gaps 3;

QY 1 MALRPSDNVYLLPSPSARVYVNDVYTRTSIFYHAGSSRLTVGNPYRVPAGGK 59  
 Db 27 MALRPSDNVYLLPSPSARVYVNDVYTRTSIFYHAGSSRLTVGNPYRVPAGGK 86  
 QY 60 QDIPKSAVQYRFRVQLPDPNKGFLPDSINYPETORLVMACAGVEIGRQPLVLSG 119  
 Db 87 QAVKSAVQYRFRVQLPDPNKGFLPDSINYPETORLVMACAGVEIGRQPLVLSG 146  
 QY 120 HPEFNKLDLTSSHAATSNVSEVDYRDNVSDYKOTQCLICGAPAIGEHMAKGTACKR 179  
 Db 147 HPEFNKLDLTSSHAATSNVSEVDYRDNVSDYKOTQCLICGAPAIGEHMAKGTACKR 206  
 QY 180 LSGDCCPLELKNITVLEGGDMVDYGYGAMDESLDTDFCEVPLDTCOSICKYPTLQMSA 239  
 Db 207 LQPDCCPLELKNITVLEGGDMVDYGYGAMDESLDTDFCEVPLDTCOSICKYPTLQMSA 266  
 QY 240 DPYDSMEFCLRRQLFAHFHFNWAGTMDVPSLYIKGT -GNRASPGSCVSPSPSG 297  
 Db 267 DPYDSMEFCLRRQLFAHFHFNWAGTMDVPSLYIKGT -GNRASPGSCVSPSPSG 326  
 QY 298 STVSDSGLFKPFWLKAQGHNGICWHNQLFTVYVDTTRSTNLTICASTQSPVPGQYD 357  
 Db 327 STVSDSGLFKPFWLKAQGHNGICWHNQLFTVYVDTTRSTNLTICASTQSPVPGQYD 386  
 QY 358 ATKFKOYSRHVEEDLOFIFOLCTITLTADVMSYIHSNNSIILEDVNGVPPPTSLVD 417  
 Db 387 PTKFKOYSRHVEEDLOFIFOLCTITLTADVMSYIHSNNSIILEDVNGVPPPTSLVD 446  
 QY 418 TYRFVQSAVATCQDAAPAEKNDPYDKLFKFNVDLKEKFSIDLDOYPLGRKFLVQAGLR 477  
 Db 447 TYRFVQSAVATCQDAAPAEKNDPYDKLFKFNVDLKEKFSIDLDOYPLGRKFLVQAGLR 506  
 QY 478 KPTIGPKRRSAP---ATTSSKPAKRVVARK 507  
 Db 507 RPTIGPKRRSAPASTSTASTASPAKRVVARK 539

## RESULT 3

P1M189

LI protein - human papillomavirus type ME180 (provirus)

C:Species: human papillomavirus type ME180

A:Note: host Homo sapiens (man)

C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 28-Jul-2000

C:Accession: B40509

R:Reuter, S.; Deilus, H.; Kahn, T.; Hofmann, B.; zur Hausen, H.; Schwarz, E.

J. Virol. 65, 5564-5568, 1991

A:Title: Characterization of a novel human papillomavirus DNA in the cervical carcinoma

A:Reference number: A40509; MUID:91374616; PMID:1716694

A:Accession: B40509

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-505 &lt;RED&gt;

A:Cross-references: GB:M73258; NID:9184383; PIDN:AAFL1010.1; PID:96478870

A:Superfamily: Papillomavirus LI protein

C:Keywords: late protein

Query Match 76.9%; Score 2155; DB 1; Length 505;

Best Local Similarity 76.7%; Pred. No. 2.1e-166;

Matches 386; Conservative 52; Mismatches 65; Indels 0; Gaps 0;

QY 1 MALRPSDNVYLLPSPSARVYVNDVYTRTSIFYHAGSSRLTVGNPYRVPAGGK 60  
 Db 1 MALRPSDNVYLLPSPSARVYVNDVYTRTSIFYHAGSSRLTVGNPYRVPAGGK 60  
 QY 61 DIPKSAVQYRFRVQLPDPNKGFLPDSINYPETORLVMACAGVEIGRQPLVLSG 120  
 Db 61 DIPKSAVQYRFRVQLPDPNKGFLPDSINYPETORLVMACAGVEIGRQPLVLSG 120  
 QY 121 PFKNKLDLTSSHAATSNVSEVDYRDNVSDYKOTQCLICGAPAIGEHMAKGTACKR 180  
 Db 121 PFKNKLDLTSSHAATSNVSEVDYRDNVSDYKOTQCLICGAPAIGEHMAKGTACKR 180  
 QY 181 SQGDCPPELKNITVLEGGDMVDYGYGAMDESLDTDFCEVPLDTCOSICKYPTLQMSA 240  
 Db 181 SQGDCPPELKNITVLEGGDMVDYGYGAMDESLDTDFCEVPLDTCOSICKYPTLQMSA 240  
 QY 241 PYGDSMEFCLRRQLFAHFHFNWAGTMDVPSLYIKGTGNRASPGSCVSPSPSG 300  
 Db 241 PYGDSMEFCLRRQLFAHFHFNWAGTMDVPSLYIKGTGNRASPGSCVSPSPSG 300  
 QY 301 TSQSGLFKPFWLKAQGHNGICWHNQLFTVYVDTTRSTNLTICASTQSPVPGQYD 360  
 Db 301 TSQSGLFKPFWLKAQGHNGICWHNQLFTVYVDTTRSTNLTICASTQSPVPGQYD 360  
 QY 361 FKQYSRHVEEDLOFIFOLCTITLTADVMSYIHSNNSIILEDVNGVPPPTSLVD 420  
 Db 361 FKQYSRHVEEDLOFIFOLCTITLTADVMSYIHSNNSIILEDVNGVPPPTSLVD 420  
 QY 421 FVQSAVATCQDAAPAEKNDPYDKLFKFNVDLKEKFSIDLDOYPLGRKFLVQAGLR 480  
 Db 421 FVQSAVATCQDAAPAEKNDPYDKLFKFNVDLKEKFSIDLDOYPLGRKFLVQAGLR 480  
 QY 481 IGPKRKRSAPATSSSKPAKRVVARK 503  
 Db 481 IGPKRKRSAPATSSSKPAKRVVARK 503

## RESULT 4

P1M139

LI protein - human papillomavirus type 39

C:Species: human papillomavirus type 39

A:Note: host Homo sapiens (man)

C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 16-Jul-1999

C:Accession: B38502

R:Voijers, C.; Strebeck, R.E.

Virology 181, 419-423, 1991

A:Title: Genome organization and nucleotide sequence of human papillomavirus type 39.

A:Reference number: A38502; MUID:91135017; PMID:1847266

A:Accession: B38502

A:Status: translation not shown

A:Molecule type: DNA  
 A:Residues: 1-505 <VOL>  
 A:Cross-references: GB:M62849; EMBL:M38185; NID:g333245; PIDN:AAA47056.1; PID:g463192  
 C:Superfamily: papillomavirus L1 protein  
 C:Keywords: late protein

Query Match 78.7%; Score 2151; DB 1; Length 505;  
 Best Local Similarity 77.3%; Pred. No. 4,3e-166;

Matches 391; Conservative 48; Mismatches 61; Indels 6; Gaps 3;

```

QY 1 MALMRSDDNTVYLPPPSVAVVNTDDYVTRTSIFTHAGSSRLITVGNPFYRVAGGANKQ 60
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1 MALMRSDDNTVYLPPPSVAVVNTDDYVTRTSIFTHAGSSRLITVGNPFYRVAGGANKQ 60
QY 61 DIPKVSAYQYRVFRVQLPDPNKGFLPDNSITYPEORLWACAGVEIGRGQPLGVLSGH 120
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 61 DIPKVSAYQYRVFRVQLPDPNKGFLPDNSITYPEORLWACAGVEIGRGQPLGVLSGH 120
QY 121 PFYNKLDLDTRESSHAATSNVSEYVDVNDVYKOTQCTICGAPALGEHNAKGTACKSRPL 180
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 121 PFYNKLDLDTRESSHAATSNVSEYVDVNDVYKOTQCTICGAPALGEHNAKGTACKSRPL 179
QY 181 SOGDCEPLELKNVLEEDGDMVDVGYGAMPFSTLQDTRKEVPLDTCOSICKPYPLQMSAD 240
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 181 SOGDCEPLELKNVLEEDGDMVDVGYGAMPFSTLQDTRKEVPLDTCOSICKPYPLQMSAD 239
QY 241 PYGDSMEFCLRRQQLFARHFNWAGTMDVPSQSLYIKGTGMASPGSCVYSPSGSIY 300
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 241 PYGDSMEFCLRRQQLFARHFNWAGTMDVPSQSLYIKGTGMASPGSCVYSPSGSIY 299
QY 301 TTSOLFENKPYWLHKAQGHNGICMHNQLFVYVDTTRSTNLTICASTGSPVPGQDANK 360
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 301 TTSOLFENKPYWLHKAQGHNGICMHNQLFVYVDTTRSTNLTICASTGSPVPGQDANK 359
QY 361 FKQYSHREYEDLQFIQCTITLFDVMSYIHSNMSILEDNFGVPPPTSLVDYTR 420
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 361 FKQYSHREYEDLQFIQCTITLFDVMSYIHSNMSILEDNFGVPPPTSLVDYTR 419
QY 421 FVQSVATTCOKDAAPAEKNDPYDKLFKFWNDLKEKESLDDQYPLGRKFLVQAGLRKPT 480
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 421 FVQSAITTCOKDAPEKNDPYDKLFKFWNDLKEKESLDDQYPLGRKFLVQAGLRKPT 479
QY 481 IGPRKRSAPSATSSKPA---KRVV 503
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 481 IGPRKR--PAASTSSATKHKRRV 503

RESULT 5
S36554
L1 protein (alternative) - human papillomavirus type 3
C:Species: human papillomavirus type 3
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36554
R:Delius, H.; Hofmann, B.
A:Description: Primer directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36554
A:Molecule type: DNA
A:Residues: 1-532 <DEL>
A:Cross-references: EMBL:X74462; NID:g397005; PIDN:CAA52474.1; PID:g397012
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

Query Match 70.2%; Score 1917.5; DB 2; Length 532;
Best Local Similarity 68.8%; Pred. No. 3.7e-147;
Matches 351; Conservative 65; Mismatches 85; Indels 9; Gaps 4;
QY 1 MALMRSDDNTVYLPPPSVAVVNTDDYVTRTSIFTHAGSSRLITVGNPFYRVAGGANKQ 60
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 29 MALMRSDDNTVYLPPPSVAVVNTDDYVTRTSIFTHAGSSRLITVGNPFYRVAGGANKQ 88
QY 61 DIPKVSAYQYRVFRVQLPDPNKGFLPDNSITYPEORLWACAGVEIGRGQPLGVLSGH 120
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 61 DIPKVSAYQYRVFRVQLPDPNKGFLPDNSITYPEORLWACAGVEIGRGQPLGVLSGH 120
QY 121 PFYNKLDLDTRESSHAATSNVSEYVDVNDVYKOTQCTICGAPALGEHNAKGTACKSRPL 180
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 121 PFYNKLDLDTRESSHAATSNVSEYVDVNDVYKOTQCTICGAPALGEHNAKGTACKSRPL 180
QY 181 SOGDCEPLELKNVLEEDGDMVDVGYGAMPFSTLQDTRKEVPLDTCOSICKPYPLQMSAD 240
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 181 SOGDCEPLELKNVLEEDGDMVDVGYGAMPFSTLQDTRKEVPLDTCOSICKPYPLQMSAD 240
QY 241 PYGDSMEFCLRRQQLFARHFNWAGTMDVPSQSLYIKGTGMASPGSCVYSPSGSIY 300
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 241 PYGDSMEFCLRRQQLFARHFNWAGTMDVPSQSLYIKGTGMASPGSCVYSPSGSIY 300

```

```

Db 89 DIPKVSAYQYRVFRVQLPDPNKGFLPDNSITYPEORLWACAGVEIGRGQPLGVLSGH 148
QY 121 PFYNKLDLDTRESSHAATSNVSEYVDVNDVYKOTQCTICGAPALGEHNAKGTACKSRPL 180
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 149 PFYNKLDLDTRESSHAATSNVSEYVDVNDVYKOTQCTICGAPALGEHNAKGTACKSRPL 207
QY 181 SOGDCEPLELKNVLEEDGDMVDVGYGAMPFSTLQDTRKEVPLDTCOSICKPYPLQMSAD 240
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 208 SPGDCEPLELITAPIDGDMVDVGYGAMPFSTLQDTRKEVPLDTCOSICKPYPLQMSAD 267
QY 241 PYGDSMEFCLRRQQLFARHFNWAGTMDVPSQSLYIKGTGMASPGSCVYSPSGSIY 300
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 241 PYGDSMEFCLRRQQLFARHFNWAGTMDVPSQSLYIKGTGMASPGSCVYSPSGSIY 300
QY 298 SIYTSOQLFNKPYWLHKAQGHNGICMHNQLFVYVDTTRSTNLTICASTGSPVPGQDANK 357
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 328 SMVTSEQLFNKPYWLHKAQGHNGICMHNQLFVYVDTTRSTNLTICASTGSPVPGQDANK 385
QY 358 ATKFKQYSHREYEDLQFIQCTITLFDVMSYIHSNMSILEDNFGVPPPTSLVDYTR 417
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 386 ATKFKQYSHREYEDLQFIQCTITLFDVMSYIHSNMSILEDNFGVPPPTSLVDYTR 445
QY 418 TYPFQSVATTCOKDAAPAEKNDPYDKLFKFWNDLKEKESLDDQYPLGRKFLVQAGLRK 477
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 446 TYPFQSVATTCOKDAAPAEKNDPYDKLFKFWNDLKEKESLDDQYPLGRKFLVQAGLRK 505
QY 478 KPTIGPRKRSAPSATSSKPAKRVRRAR 507
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 506 RSSISVAKR--SATYTSRTAAAKRKTKK 532

RESULT 6
S36537
L1 protein - human papillomavirus type 10
C:Species: human papillomavirus type 10
C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36537
R:Delius, H.; Hofmann, B.
A:Description: Primer directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36537
A:Molecule type: DNA
A:Residues: 1-531 <DEL>
A:Cross-references: EMBL:X74465; NID:g396901; PIDN:CAA52494.1; PID:g396908
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

Query Match 69.8%; Score 1909; DB 2; Length 531;
Best Local Similarity 67.9%; Pred. No. 1.8e-146;
Matches 342; Conservative 70; Mismatches 90; Indels 2; Gaps 2;
QY 1 MALMRSDDNTVYLPPPSVAVVNTDDYVTRTSIFTHAGSSRLITVGNPFYRVAGGANKQ 60
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 29 MALMRSDDNTVYLPPPSVAVVNTDDYVTRTSIFTHAGSSRLITVGNPFYRVAGGANKQ 88
QY 61 DIPKVSAYQYRVFRVQLPDPNKGFLPDNSITYPEORLWACAGVEIGRGQPLGVLSGH 120
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 61 DIPKVSAYQYRVFRVQLPDPNKGFLPDNSITYPEORLWACAGVEIGRGQPLGVLSGH 148
QY 121 PFYNKLDLDTRESSHAATSNVSEYVDVNDVYKOTQCTICGAPALGEHNAKGTACKSRPL 180
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 121 PFYNKLDLDTRESSHAATSNVSEYVDVNDVYKOTQCTICGAPALGEHNAKGTACKSRPL 180
QY 149 PFYNKLDLDTRESSHAATSNVSEYVDVNDVYKOTQCTICGAPALGEHNAKGTACKSRPL 207
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 181 SOGDCEPLELKNVLEEDGDMVDVGYGAMPFSTLQDTRKEVPLDTCOSICKPYPLQMSAD 240
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 208 SPGDCEPLELITAPIDGDMVDVGYGAMPFSTLQDTRKEVPLDTCOSICKPYPLQMSAD 267
QY 241 PYGDSMEFCLRRQQLFARHFNWAGTMDVPSQSLYIKGTGMASPGSCVYSPSGSIY 300
   ||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 241 PYGDSMEFCLRRQQLFARHFNWAGTMDVPSQSLYIKGTGMASPGSCVYSPSGSIY 300

```

```
OY 301 TSDSOLFENKPYWLAKAOGHNNICWNNOLFVTVVDTTRSTNLITICASTQSPGQYDATK 360
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 328 TSEOLFENKPYWLAKAOGHNNICWNNOLFVTVVDTTRSTNLITICASTQSPGQYDATK 387
OY 361 FKQSRHVEEYDLOFIFOLCTITLTADVMYSIHSMNSIILEDMMNGVPPPTSLVDYTR 420
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 388 FKEIRHVEEYDLOFIFOLCTITLTADVMYSIHSMNSIILEDMMNGVPPPTSLVDYTR 447
OY 421 FVQSVATTCQKDAAPAEKNDPYDKLFKNVNDLKEKFSIDLDOYPLGRKFLVOAGLRKPT 480
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 448 FLSSAATCQKDPTEKQDYAKLNFWDVDLKDREFSLDLSQPLGRKFLQLQGLVRSNA 507
OY 481 IGRKRSAPSATSSKPAKRRVR 504
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 508 VSVRRKRSAPSA-TGSTAKRRRTK 530
```

## RESULT 7

```
S36508
L1 protein - human papillomavirus type 30
C:Species: human papillomavirus type 30
C:Date: 20-Feb-1995 #sequence_rev1sion 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36508
R:Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36508
A:Molecule type: DNA
A:Residues: 1-508 <DEL>
A:Cross-references: EMBL:X74474; NID:q396973; PIDN:CAA52548.1; PID:q396980
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein
```

```
Query Match 68.8%; Score 1879; DB 2; Length 508;
Best Local Similarity 67.1%; Pred. No. 4.5e-144;
Matches 337; Conservative 82; Mismatches 81; Indels 2; Gaps 1;
```

```
OY 1 MALRPSDNITYLLPPPSYARVYVNDYVTRTSIFTHAGSSRLTYGAFYFVPPAGGKNO 60
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 8 MAVMPSSTKYLLPPTPVSKYVPTDAYKRNIFYHAGSSRLTAVGHYYSISRKNSKT 67
OY 61 DIPKSAVQYVRFVQLPDPNKFGLPNSIYNPETORLWMAAGVEIGRGOPLGVLGSGH 120
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 68 DVPKSAVQYVRFVQLPDPNKFGLPNSIYNPETORLWMAAGVEIGRGOPLGVLGSGH 127
OY 121 PFYKLDDESSHAATSVSEVDVNDVYKQFQOLICGAPALGEHMAKGTACKSRPL 180
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 128 PLFNKLDDESSHAATSVSEVDVNDVYKQFQOLICGAPALGEHMAKGTACKSRAP 187
OY 181 SOGCPPELEKNTYLEGDMDVDTGFGAMDPESTLDTCFVPLDTCOSICKYPTDLOKMSAD 240
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 188 AOGDCPPELEKNTYLEGDMDVDTGFGAMDPESTLDTCFVPLDTCOSICKYPTDLOKMSAD 247
OY 241 PYGDSMFECLEKNTYLEGDMDVDTGFGAMDPESTLDTCFVPLDTCOSICKYPTDLOKMSAD 300
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 248 AYGDMSMAYLREDFLFAHFNNRAGTMDVTPQSLYIKGTGMKSPGSCVYSPSPSGSIV 307
OY 301 TSDSOLFENKPYWLAKAOGHNNICWNNOLFVTVVDTTRSTNLITICASTQSPGQYDATK 360
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 308 TSEOLFENKPYWLAKAOGHNNICWNNOLFVTVVDTTRSTNLITICASTQSPGQYDATK 365
OY 361 FKQSRHVEEYDLOFIFOLCTITLTADVMYSIHSMNSIILEDMMNGVPPPTSLVDYTR 420
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 366 IKQYRHEEYELDFVQLCKISLAEETMAYLHTMNSLITGWMNGSLPPATSLDEDYR 425
OY 421 FVQSVATTCQKDAAPAEKNDPYDKLFKNVNDLKEKFSIDLDOYPLGRKFLVOAGLRKPT 480
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 426 YKSLATTCQKDPTEKQDYAKLNFWDVDLKDREFSLDLSQPLGRKFLQLQGLVRSNA 485
OY 481 IGRKRSAPSATSSKPAKRRVR 502
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 486 TTTKRSAPSSSTSTSPSAKRR 507
```

## RESULT 8

S36549

```
L1 protein - human papillomavirus type 26
C:Species: human papillomavirus type 26
C:Date: 20-Feb-1995 #sequence_rev1sion 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36549
R:Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36549
A:Molecule type: DNA
A:Residues: 1-503 <DEL>
A:Cross-references: EMBL:X74472; NID:q396956; PIDN:CAA52535.1; PID:q396963
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein
```

```
Query Match 68.1%; Score 1861.5; DB 2; Length 503;
Best Local Similarity 67.3%; Pred. No. 1.2e-142;
Matches 337; Conservative 70; Mismatches 91; Indels 3; Gaps 3;
```

```
OY 1 MALRPSDNITYLLPPPSYARVYVNDYVTRTSIFTHAGSSRLTYGAFYFVPPAGGKNO 60
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MALRPSDNITYLLPPPSYARVYVNDYVTRTSIFTHAGSSRLTYGAFYFVPPAGGKNO 60
OY 61 DIPKSAVQYVRFVQLPDPNKFGLPNSIYNPETORLWMAAGVEIGRGOPLGVLGSGH 120
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 60 DIPKSAVQYVRFVQLPDPNKFGLPNSIYNPETORLWMAAGVEIGRGOPLGVLGSGH 119
OY 121 PFYKLDDESSHAATSVSEVDVNDVYKQFQOLICGAPALGEHMAKGTACKSRPL 180
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 PLFNKLDDESSHAATSVSEVDVNDVYKQFQOLICGAPALGEHMAKGTACKSRPL 179
OY 181 SOGCPPELEKNTYLEGDMDVDTGFGAMDPESTLDTCFVPLDTCOSICKYPTDLOKMSAD 240
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 180 OQDCPPELEKNTYLEGDMDVDTGFGAMDPESTLDTCFVPLDTCOSICKYPTDLOKMSAD 239
OY 241 PYGDSMFECLEKNTYLEGDMDVDTGFGAMDPESTLDTCFVPLDTCOSICKYPTDLOKMSAD 300
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 240 TYGDSMFECLEKNTYLEGDMDVDTGFGAMDPESTLDTCFVPLDTCOSICKYPTDLOKMSAD 299
OY 300 VTSSOLFENKPYWLAKAOGHNNICWNNOLFVTVVDTTRSTNLITICASTQSPGQYDATK 359
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 300 VTSSOLFENKPYWLAKAOGHNNICWNNOLFVTVVDTTRSTNLITICASTQSPGQYDATK 359
OY 360 FKQSRHVEEYDLOFIFOLCTITLTADVMYSIHSMNSIILEDMMNGVPPPTSLVDYTR 419
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 360 DYKQFIRHGEYELDFVQLCKITLTADVMYSIHSMNSIILEDMMNGVPPPTSLVDYTR 419
OY 420 FVQSVATTCQKDAAPAEKNDPYDKLFKNVNDLKEKFSIDLDOYPLGRKFLVOAGLRKPT 479
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 420 RFIKNSATTCQKDAAPAEKNDPYDKLFKNVNDLKEKFSIDLDOYPLGRKFLVOAGLRKPT 479
OY 480 TIGRKRSPAPSATSSKPAKRRVR 500
    ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 480 KLG-TKRP LSTSSSTRKRR 499
```

## RESULT 9

S36531

```
L1 protein - human papillomavirus type 53
C:Species: human papillomavirus type 53
C:Date: 20-Feb-1995 #sequence_rev1sion 20-Feb-1995 #text_change 26-Aug-1999
C:Accession: S36531
R:Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36531
A:Molecule type: DNA
A:Residues: 1-499 <DEL>
```



```

Db      121 HPLNKFDDTENNENRAGGPGTDNRRCISMDYKOTQLCLGCKPRLGEHMGKSPCSNNA 180
      180 LSGDDPPLELKTWVLEDDGMVDTGYGAMDFSTLDQKCEVPLDICOISICKYDYLOMSA 239
      181 ITGDDPPLELKTWVLEDDGMVDTGYGAMDFSTLDQKCEVPLDICOISICKYDYLOMSA 240
      240 DPYGSDFECLRRQQLFARHFNWAGTMDTPQSLYKGTGMRASPGSCVSPSPSGST 299
      241 EPGDGLTFEFLRREGMEVFRHFNRSRGVGSVPTDLYIKSGSGATLANSTYPTPSGSM 300
      300 VTSDSOLFENKPYWLHRAQGNNGICWNOLFTVTVDTTRSTNLTICASTQSPVPGQYDAT 359
      301 VTSDAOLFENKPYWMORAGNNGICWNOLFTVTVDTTRSTNLTICASTQSPVPGQYDAT 359
      360 KFKOYSHVBEYDLOLTFOLCTITLTADVMSYIHSNNSLTEDMNGVPPPTSLVDYR 419
      360 NFKYELHGEEDLOLTFOLCTITLTADVMSYIHSNNSLTEDMNGVPPPTSLVDYR 419
      420 RPYOVALTCQKDAAPENKDPYDKLTFMNVDLKEKESLDLDQYPLGRKFLVQAGLRKP 479
      420 RPYOVALTCQKDAAPENKDPYDKLTFMNVDLKEKESLDLDQYPLGRKFLVQAGLRKP 479
      480 TIGPKRSAPSATSSKPAKRVYR 504
      480 KFKRGRSAPSASTTT-PAKRKRK 503

```

## RESULT 12

```

S36526
L1 protein - human papillomavirus type 35H
C:Species: human papillomavirus type 35H
C>Date: 09-Dec-1993 #sequence_revision 26-Jul-1996 #text_change 26-Aug-1999
C:Accession: S36526
R:Delius, H.; Hofmann, B.
submitted to the EMBL Data Library, August 1993
A:Description: Primer-directed sequencing of human papillomavirus types.
A:Reference number: S36469
A:Accession: S36526
A:Molecule type: DNA
A:Residues: 1-502 <DEL>
A:Cross-references: EMBL:X74477; NID:g396997; PIDN:CAA52566.1; PID:g397004
A:Experimental source: strain 35H
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

```

```

Query Match      66.6%; Score 1820.5; DB 2; Length 502;
Best Local Similarity 65.1%; Pred. No. 2.4e-139;
Matches 329; Conservative 70; Mismatches 103; Indels 3; Gaps 2;

QY      1 MALWRPSDNTVYLPSPSVARVNTDVTYRTSIFHYAGSSRLTLVGNPFRVAGGKNQ 60
      1 MSLMRSSEATVYLPSPSVARVNTDVTYRTSIFHYAGSSRLTLVGNPFRVAGGKNQ 60
      61 DIRKVSAYQYRVRVOLPDPNKGFLPDNSIYNPETORLVWACAGVEIGRQPLGVLSGH 120
      61 AVKRVSGIQYRVRVOLPDPNKGFLPDNSIYNPETORLVWACAGVEIGRQPLGVLSGH 120
      121 PFYNNKLDDESSHAATSNVEDVRYDQYKOTQLCTICAPRIGHMAKGTACKSRPL 180
      121 PLANKLDDESSHAATSNVEDVRYDQYKOTQLCTICAPRIGHMAKGTACKSRPL 180
      181 SGGDCPPELEKNTVLEDDGMVDTGYGAMDFSTLDQKCEVPLDICOISICKYDYLOMSA 240
      181 KAGCPEPPELEKNTVLEDDGMVDTGYGAMDFSTLDQKCEVPLDICOISICKYDYLOMSA 240
      241 PYGDSMFECRLRRQQLFARHFNWAGTMDTPQSLYKGTGMRASPGSCVSPSPSGSIY 300
      241 PYGDSMFECRLRRQQLFARHFNWAGTMDTPQSLYKGTGMRASPGSCVSPSPSGSIY 300
      301 TSDSOLFENKPYWLHRAQGNNGICWNOLFTVTVDTTRSTNLTICASTQSPVPGQYDAT 360
      301 TSDSOLFENKPYWLHRAQGNNGICWNOLFTVTVDTTRSTNLTICASTQSPVPGQYDAT 360
      299 TSDSOLFENKPYWLHRAQGNNGICWNOLFTVTVDTTRSTNLTICASTQSPVPGQYDAT 357

```

```

QY      361 KFKOYSHVBEYDLOLTFOLCTITLTADVMSYIHSNNSLTEDMNGVPPPTSLVDYR 420
      358 KFKYELHGEEDLOLTFOLCTITLTADVMSYIHSNNSLTEDMNGVPPPTSLVDYR 417
      421 FVOSVALTCQKDAAPENKDPYDKLTFMNVDLKEKESLDLDQYPLGRKFLVQAGLRKP 480
      418 YVTSQAVTCCQKDAAPENKDPYDKLTFMNVDLKEKESLDLDQYPLGRKFLVQAGLRKP 477
      481 IGPKRSAPSATSSKPAKRVYR 505
      478 FRLKRAAPASATSSKPAKRVYR 502

```

## RESULT 13

```

L1 protein - human papillomavirus type 51
C:Species: human papillomavirus type 51
A:Note: host Homo sapiens (man)
C>Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 16-Jul-1999
C:Accession: G40415; G44889
R:lungu, O.; Crum, C.P.; Silverstein, S.J.
J. Virol. 65, 4216-4225, 1991
A:Title: Biologic properties and nucleotide sequence analysis of human papillomavirus
A:Reference number: A40415; M01D:91303675; PMID:1649326
A:Accession: G40415
A:Status: translation not shown
A:Molecule type: DNA
A:Residues: 1-504 <LUN>
A:Cross-references: GB:M62877
R:van den Brule, A.J.; Snijders, P.J.; Raaphorst, P.M.; Schrijnemakers, H.F.; Delius,
J. Clin. Microbiol. 30, 1716-1721, 1992
A:Title: General primer polymerase chain reaction in combination with sequence analysis
A:Reference number: A44889; M01D:92332706; PMID:1321168
A:Accession: G44889
A:Molecule type: DNA
A:Residues: 337-369 <VAN>
A:Cross-references: GB:S40272; NID:g251694; PIDN:AAB22568.1; PID:g251695
A:Experimental source: mucosotropic type 51, cervical smear
A:Note: sequence extracted from NCBI backbone (NCBIN:109397, NCBI:P.109409)
C:Superfamily: papillomavirus L1 protein
C:Keywords: late protein

```

```

Query Match      66.0%; Score 1804.5; DB 1; Length 504;
Best Local Similarity 64.0%; Pred. No. 4.8e-138;
Matches 323; Conservative 83; Mismatches 96; Indels 3; Gaps 3;

QY      1 MALWRPSDNTVYLPSPSVARVNTDVTYRTSIFHYAGSSRLTLVGNPFRVAGGKNQ 60
      1 MALWRPSDNTVYLPSPSVARVNTDVTYRTSIFHYAGSSRLTLVGNPFRVAGGKNQ 60
      61 DIRKVSAYQYRVRVOLPDPNKGFLPDNSIYNPETORLVWACAGVEIGRQPLGVLSGH 120
      61 DIRKVSAYQYRVRVOLPDPNKGFLPDNSIYNPETORLVWACAGVEIGRQPLGVLSGH 120
      121 PFYNNKLDDESSHAATSNVEDVRYDQYKOTQLCTICAPRIGHMAKGTACKSRPL 180
      121 PFYNNKLDDESSHAATSNVEDVRYDQYKOTQLCTICAPRIGHMAKGTACKSRPL 180
      181 SGGDCPPELEKNTVLEDDGMVDTGYGAMDFSTLDQKCEVPLDICOISICKYDYLOMSA 240
      181 SGGDCPPELEKNTVLEDDGMVDTGYGAMDFSTLDQKCEVPLDICOISICKYDYLOMSA 240
      241 PYGDSMFECRLRRQQLFARHFNWAGTMDTPQSLYKGTGMRASPGSCVSPSPSGSIY 300
      241 PYGDSMFECRLRRQQLFARHFNWAGTMDTPQSLYKGTGMRASPGSCVSPSPSGSIY 300
      301 TSDSOLFENKPYWLHRAQGNNGICWNOLFTVTVDTTRSTNLTICASTQSPVPGQYDAT 360
      301 TSDSOLFENKPYWLHRAQGNNGICWNOLFTVTVDTTRSTNLTICASTQSPVPGQYDAT 360
      360 KFKOYSHVBEYDLOLTFOLCTITLTADVMSYIHSNNSLTEDMNGVPPPTSLVDYR 419

```

Db 359 NRQYIRHGEVELOPIFOLCKITLTTEYMAILHTMDPTLLEQMNGLLPPSASILEDAY 418  
 QY 420 REVQSVATTCORDAAPAEKNKDPYDKLFENNVLDLKEKFSLDLQYPLGRFFLYOAGLRKP 479  
 Db 419 RRVNMAATSCQKQDTPPOAKDPDLAKYKFDVDLKEKFSLDLQFALGRFFLYOAGLRKP 478  
 QY 480 TIGPKRRSAPSAATSSKPAKRVVR 504  
 Db 479 RPKLRKPASSASSSSSSAKRRRVK 503

## RESULT 14

PLWLS8  
 L1 protein - human papillomavirus type 58  
 C:Species: human papillomavirus type 58  
 A:Note: host Homo sapiens (man)  
 C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 21-Jul-2000  
 C:Accession: G36779  
 R:Kiriil, Y.; Iwamoto, S.; Matsukura, T.  
 Virology 185, 424-427, 1991  
 A:Title: Human papillomavirus type 58 DNA sequence.  
 A:Reference number: A36779; MID:92024102; PMID:1656594  
 A:Accession: G36779  
 A:Status: translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-524 <KIR>  
 A:Cross-references: GB:D90400; MID:9222386; PIDN:BAA31851.1; PID:9337104  
 C:Superfamily: papillomavirus L1 protein  
 C:Keywords: late protein

Query Match 66.0%; Score 1804.5; DB 1; Length 524;  
 Best Local Similarity 65.0%; Pred. No. 5,1e-138;  
 Matches 327; Conservative 73; Mismatches 96; Indels 7; Gaps 4;

QY 1 MALMRSDNTVYLPPPSVARVYNTDDYVRTSTFYHAGSSRLTLVGNPFYRV--PAGGG 60  
 Db 27 MSWMRSEATVYLPVPVSKVYSTDEYVRSITYYAGSSRLLVGNPFYSIKSPNNKK 86  
 QY 61 D-IPRVSAYQYVFRVQLDPDNKFGIPDNSTYNETQRLVWACAGVEIGRQPLGVLSG 119  
 Db 87 YVAPKVSGLQYVFRVRLDPDNKFGIPDNSTYNETQRLVWACAGVEIGRQPLGVLSG 146  
 QY 120 HEPYKRLDPTSSHAATSNVSEVDYRNVDRKOTQCLIGCAPAIGEMAKTACKSRP 179  
 Db 147 HFLYLNKFDDETSTNRYPAQPGSDNRECLSMQYKQQLCIGCKPPTGEMGKGVACNNNA 206  
 QY 180 LSGQDCPPLKNTVLEDDMYDTGYGAMDFSTLDTKCEVPLDQICQISCKYPDYIQMSA 239  
 Db 207 AAT-DQPLFLFNSTIEDMDYDTGRCMDFTGLQANKSDVPLDQICNSTCKYPDYIKMAS 265  
 QY 240 DRYGSMFECLEARQOLFAHFHFNRAAGTMDVTPQSLYIKGTGNRASPGSCVYSPSPSGSI 299  
 Db 266 EBYGSLFEFFLRQDMFVHFHFNRAAGTMDVTPQSLYIKGTGNRASPGSCVYSPSPSGSI 325  
 QY 300 VNSDSOLFNRKPYMLHKAQGHNNIGCHNOLFVYVDTTRSTNLTICASTQSPVPGYDMF 359  
 Db 326 VTSESQLFKPYMLDRAQGHNNIGCHNOLFVYVDTTRSTNLTIC--TEVTEGKYKND 383  
 QY 360 KKKQYSRHVEYDLOFIFOLCTITTLADVMSYIHSNNSILLEDNMGVPPPTTSLVDTY 419  
 Db 384 NKKEYVRHVEYDLOFVPOCLKITLTAELMTYIHRMDSNILDMDQGLPPPSASIQDYY 443  
 QY 420 REVQSVATTCORDAAPAEKNKDPYDKLFENNVLDLKEKFSLDLQYPLGRFFLYOAGLRKP 479  
 Db 444 RRVNMAATSCQKQDTPPOAKDPDLAKYKFDVDLKEKFSLDLQFALGRFFLYOAGLRKP 503  
 QY 480 TIGPKRRSAPSAATSSKPAKRVVR 502  
 Db 504 RL---KRSAPITRAPSTAKKKV 523

RESULT 15  
 S36578

L1 protein - human papillomavirus type 52  
 C:Species: human papillomavirus type 52  
 C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999  
 C:Accession: S36578; MID:92332706; PMID:1321168  
 R:Deilus, H.; Hofmann, B.  
 Submitted to the EMBL Data Library, August 1993  
 A:Description: Primer-directed sequencing of human papillomavirus types.  
 A:Reference number: S36469  
 A:Accession: S36578  
 A:Molecule type: DNA  
 A:Residues: 1-529 <DEL>

A:Cross-references: EMBL:X74481; MID:9397038; PIDN:CAA52590.1; PID:9397045  
 R:van den Brule, A.J.; Snijders, P.J.; Raaphorst, P.M.; Schilfjennemakers, H.F.; Deilus, J. Clin. Microbiol. 30, 1716-1721, 1992  
 A:Title: General primer polymerase chain reaction in combination with sequence analysis  
 A:Reference number: A44889; MID:92332706; PMID:1321168  
 A:Accession: I44889  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 368-399 <VAN>  
 A:Cross-references: GB:S40277; MID:9251698; PIDN:AAB22570.1; PID:9251699  
 A:Experimental source: type Xa, cervical carcinoma in situ  
 A:Note: sequence extracted from NCBI backbone (NCBI:109399, NCBI:109447)  
 C:Superfamily: papillomavirus L1 protein  
 C:Keywords: late protein

Query Match 65.9%; Score 1801.5; DB 2; Length 529;  
 Best Local Similarity 65.0%; Pred. No. 9e-138;  
 Matches 328; Conservative 69; Mismatches 101; Indels 7; Gaps 3;

QY 1 MALMRSDNTVYLPPPSVARVYNTDDYVRTSTFYHAGSSRLTLVGNPFYRV--PAGGG 57  
 Db 27 MSWMRSEATVYLPVPVSKVYSTDEYVRSITYYAGSSRLTLVGNPFYSIKNTSSGNG-86  
 QY 58 NKQDVKVSAYQYVFRVQLDPDNKFGIPDNSTYNETQRLVWACAGVEIGRQPLGVLSG 117  
 Db 87 KVLVPRVSGLOQYVFRVRLDPDNKFGIPDNSTYNETQRLVWACAGVEIGRQPLGVLSG 146  
 QY 118 SGHFFYKRLDPTSSHAATSNVSEVDYRNVDRKOTQCLIGCAPAIGEMAKTACKSRP 177  
 Db 147 SGHPLNKFDDETSTNRYPAQPGSDNRECLSMQYKQQLCIGCKPPTGEMGKGVACNNNA 206  
 QY 178 RPLSQDCPPLKNTVLEDDMYDTGYGAMDFSTLDTKCEVPLDQICQISCKYPDYIQMSA 237  
 Db 207 NSGNPDCPPLQILINSVYDGYGAMDFSTLDTKCEVPLDQICQISCKYPDYIQMSA 266  
 QY 238 SADPYGDSMEFECLEARQOLFAHFHFNRAAGTMDVTPQSLYIKG--TGMRASPGSCVYSPSPSGSI 295  
 Db 267 ASEPYGDSLFEFFLRQDMFVHFHFNRAAGTMDVTPQSLYIKG--TGMRASPGSCVYSPSPSGSI 326  
 QY 296 SGIATVSDSOLFNRKPYMLHKAQGHNNIGCHNOLFVYVDTTRSTNLTICASTQSPVPGQ 355  
 Db 327 SGMVTVSESQLFKPYMLDRAQGHNNIGCHNOLFVYVDTTRSTNLTICAEVKK--EST 384  
 QY 356 YDATKQYSRHVEYDLOFIFOLCTITTLADVMSYIHSNNSILLEDNMGVPPPTTSLVDTY 415  
 Db 385 YKNEKREYLRHVEYDLOFIFOLCTITTLADVMSYIHSNNSILLEDNMGVPPPTTSLVDTY 444  
 QY 416 VDRYFVQSVATTCORDAAPAEKNKDPYDKLFENNVLDLKEKFSLDLQYPLGRFFLYOAGLRKP 475  
 Db 445 EDYRFRVYATTCORDAAPAEKNKDPYDKLFENNVLDLKEKFSLDLQYPLGRFFLYOAGLRKP 504  
 QY 476 RRPITGPKRRSAPSAATSSKPAKRVVR 500  
 Db 505 QARPKLRPSSAPSTSTKKKV 529

Search completed: October 17, 2003, 11:03:01  
 Job time : 29.2831 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2003, 10:55:49 ; Search time 25.7169 Seconds  
(without alignments)  
1723.913 Million cell updates/sec

Title: US-08-913-644-4

Perfect score: 2405  
Sequence: 1 MWSHRAARRRRASVTDLKYT.....FIPKRRKRVPEFADGFVAA 461

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*  
Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2317.5	96.4	462	1 P2WL18	L2 protein - human
2	1973	82.0	463	2 S36565	L2 protein - human
3	1716	71.4	469	1 P2WLPR	L2 protein - human
4	1686.5	70.1	470	1 P2WL39	L2 protein - human
5	1532.5	63.7	470	2 S36536	L2 protein - human
6	1530	63.6	463	2 S36530	L2 protein - human
7	1529	63.6	463	2 S36507	L2 protein - human
8	1513	62.9	473	2 S36553	L2 protein - human
9	1486.5	61.8	464	2 S36582	L2 protein - human
10	1483.5	61.7	472	2 S36548	L2 protein - human
11	1466.5	61.0	468	1 P2WL51	L2 protein - human
12	1457	60.6	524	1 S15619	L2 protein - human
13	1423.5	59.2	464	2 S36501	L2 protein - human
14	1387.5	57.7	465	1 S15626	L2 protein - human
15	1164.5	48.4	466	1 P2WL31	L2 protein - human
16	1158	48.1	463	1 P2WL13	L2 protein - human
17	1149.5	47.8	455	1 P2WL11	L2 protein - human
18	1138.5	47.3	466	2 P2WLRL	L2 protein - human
19	1133.5	47.1	466	2 S36577	L2 protein - human
20	1133.5	47.1	473	1 P2WLHS	L2 protein - human
21	1131	47.0	459	1 P2WL6	L2 protein - human
22	1117	46.4	477	1 P2WL42	L2 protein - human
23	1116.5	46.4	472	1 P2WL58	L2 protein - human
24	1114	46.3	467	1 P2WL33	L2 protein - human
25	1110	46.2	463	1 P2WLCL	L2 protein - human
26	1108.5	46.1	456	2 S36588	L2 protein - human
27	1088	45.2	472	2 S36519	L2 protein - human
28	1087.5	45.2	476	2 S36513	L2 protein - human
29	1069	44.4	467	2 S36559	L2 protein - human

30	1060	44.1	469	2 S36525	L2 protein - human
31	1053	43.8	469	1 P2WL35	L2 protein - human
32	642.5	26.7	519	2 S36471	L2 protein - human
33	639.5	26.6	533	2 S36594	L2 protein - human
34	628.5	26.1	533	2 S36477	L2 protein - human
35	627	26.1	492	1 P2WLRB	L2 protein - human
36	625	26.0	518	1 P2WL47	L2 protein - human
37	621	25.8	518	1 P2WL85	L2 protein - human
38	620.5	25.8	520	2 S36485	L2 protein - human
39	608.5	25.3	520	2 S36495	L2 protein - human
40	606	25.2	518	1 P2WL5	L2 protein - human
41	604.5	25.1	524	2 S36483	L2 protein - human
42	597	24.8	518	2 S36542	L2 protein - human
43	591.5	24.6	521	2 S36571	L2 protein - human
44	586	24.4	518	1 P2WL8	L2 protein - human
45	574	23.9	507	1 P2WL	L2 protein - human

ALIGNMENTS

RESULT 1

L2 protein - human papillomavirus type 18  
C/Species: human papillomavirus type 18  
C/Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 16-Jul-1999  
C/Accession: B26251  
R/Col: S.T.; Danos. O.  
J. Mol. Biol. 193, 599-608, 1987  
A/Title: Nucleotide sequence and comparative analysis of the human papillomavirus typ  
A/Reference number: A92937; MUID:87283882; PMID:3039146  
A/Accession: B26251  
A/Molecule type: DNA  
A/Residues: 1462 <COL>  
A/Cross-references: GB:X05015; NID:960975; PIDN:CAA28670.1; PID:960982  
C/Superfamily: papillomavirus L2 protein  
C/Keywords: late protein

Query Match	96.4%	Score 2317.5	DB 1:	Length 462:
Best Local Similarity	96.8%	Pred. No. 1.6e-130:		
Matches 447:	Conservative 5:	Mismatches 9:	Indels 1:	Gaps 1:
QY	1 MWSHRAARRRRASVTDLKYTKCKOSGTCPSDVVKVGGTTLADKILLOMSSLGITGLGIG	60		
DB	1 MWSHRAARRRRASVTDLKYTKCKOSGTCPPVYKVEGTTLADKILLOMSSLGITGLGIG	60		
QY	61 TSGGTGRTGYIPLGGRSNTVDVGPTRPPVIEVGPDPDSIVTLIEDSVVTSGAPRP	120		
DB	61 TSGGTGRTGYIPLGGRSNTVDVGPTRPPVIEVGPDPDSIVTLIEDSVVTSGAPRP	120		
QY	121 TFGTSGEDTTSAGTTTPAVLDTTPSSTSYSTNTFTNPAPFSDPSIEVPQGEVSGNV	180		
DB	121 TFGTSGEDTTSAGTTTPAVLDTTPSSTSYSTNTFTNPAPFSDPSIEVPQGEVAGNV	180		
QY	181 FVGTPTSGTHGYEIEIQLTFEASSGTGEEPISSPPLFTVRRVAGPRILYSRAYOOVSANPE	240		
DB	181 FVGTPTSGTHGYEIEIQLTFEASSGTGEEPISSPPLFTVRRVAGPRILYSRAYOOVSANPE	240		
QY	241 FLTRPSSLITYDNPAPFVDVTLTFEPRSNVPSDFMDIILRLRPAALTSRGTVRSRLG	300		
DB	241 FLTRPSSLITYDNPAPFVDVTLTFEPRSNVPSDFMDIILRLRPAALTSRGTVRSRLG	300		
QY	301 QRATMTFRSGTOIGAVVHFYHDSPLAPSEYIEIQLPVASATDNGLFDTIADDDPAMP	360		
DB	301 QRATMTFRSGTOIGAVVHFYHDSPLAPSEYIEIQLPVASATDNGLFDTIADDDPAMP	360		
QY	361 VPSRPTTSSAVSTYSPITSSASSYSNVTVPLTSSMDVPVYTGPDILP-PTSWPILVSP	419		
DB	361 VPSRPTTSSAVSTYSPITSSASSYSNVTVPLTSSMDVPVYTGPDILP-PTSWPILVSP	419		
QY	420 APASTGYIGIHGTHYILMPLYYFIPKRRKRVPEFADGFVAA 461			
DB	421 APASTGYIGIHGTHYILMPLYYFIPKRRKRVPEFADGFVAA 462			

RESULT 2  
S36565  
L2 protein - human papillomavirus type 45  
C:Species: human papillomavirus type 45  
C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 26-Aug-1999  
C:Accession: S36565  
R:Delius, H.; Hofmann, B.  
Submitted to the EMBL Data Library, August 1993  
A:Description: Primer-directed sequencing of human papillomavirus types.  
A:Reference number: S36469  
A:Accession: S36565  
A:Molecule type: DNA  
A:Residues: 1-463 <DEL>  
A:Cross-references: EMBL:X74479; NID:9397022; PIDN:CAA52577.1; PID:9397028  
C:Superfamily: papillomavirus L2 protein  
C:Keywords: late protein

Query Match 82.0%; Score 1973; DB 2; Length 463;  
Best Local Similarity 81.7%; Pred. No. 4,8e-110;  
Matches 379; Conservative 33; Mismatches 48; Indels 4; Gaps 3;

```

OY 1 MSHRAARRKASVTDLYKTCCKOSGCPSPDVNKEGTTLADKTLQWSSLGIFLGIG 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MSHRAARRKASVTDLYKTCCKOSGCPSPDVNKEGTTLADKTLQWSSLGIFLGIG 60

OY 61 TSGGTGRTGYIPLGGSNTVDVGPTRPVVIEPVGPTDPSITVLTIEDSSVVTSGARP 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 TSGGTGRTGYIPLGGSNTVDVGPTRPVVIEPVGPTDPSITVLTIEDSSVVTSGARP 120

OY 121 TFGTSGFDITSAGTTTTPAVLDITPSSTSVISTNTNTNPAFSPSIIIEVPQTEVSGNV 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 TFGTSGFEITSSSTTTPAVLDITPSSTSVISTNTNTNPAFSPSIIIEVPQTEVSGNI 180

OY 181 FVGTPGTGTYEIEIPQTRASSSGTEPISSPTPLPVRRVAGRLXSRAGQVRSVANPE 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 FVGTPGTGTYEIEIPQTRASSSGTEPISSPTPLPVRRVAGRLXSRAGQVRSVTSQ 240

OY 241 FLTRPSSLITYDNPAPFVDTTLTFEPRSNVPSDFMDIIRLHRPALTSRRGTVRSRLG 300
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 FLTRPSSLITYDNPAPFVDTTLTFEPRSNVPSDFMDIIRLHRPALTSRRGTVRSRLG 300

OY 301 ORATMFRSGTQIGARVHFHDISPIASPEYIELQPLVSTENGFLDYADIADIDPAMP 360
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 301 ORATMFRSGTQIGARVHFHDISPIASPEYIELQPLVSTENGFLDYADIDPAMP 360

OY 361 VPSHPTTSSAVSYSPYSS--ASSYSNVVPLTSSMDVPVYTGPDITLPP--TSVMPYVS 417
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 361 TPTIHSFTYKSLTLPSTAASSYSNVVPLTSSMDVPVYTGPDITLPP--TSVMPYVS 419

OY 418 PTPASTQYIGIHGTHYLLPPLYEIPKRRRVYFFADGFVA 461
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 420 PTPASTQYIGIHGTHYLLPPLYEIPKRRRVYFFADGFVA 463

```

## RESULT 3

P2MLPR  
L2 protein - human papillomavirus type ME180 (provirus)  
C:Species: human papillomavirus type ME180  
A:Note: host Homo sapiens (man)  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 28-Jul-2000  
C:Accession: A40509  
R:Reuter, S.; Delius, H.; Kahn, T.; Hofmann, B.; zur Hausen, H.; Schwarz, E.  
J. Virol. 65, 5564-5568, 1991  
A:Title: Characterization of a novel human papillomavirus DNA in the cervical carcinoma  
A:Reference number: A40509; MUID:91374616; PMID:1716694  
A:Accession: A40509  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-469 <REU>  
A:Cross-references: GB:M73258; NID:9184383; PIDN:AAF14009.1; PID:g6478869  
C:Superfamily: papillomavirus L2 protein

C:Keywords: late protein

Query Match 71.4%; Score 1716; DB 1; Length 469;  
Best Local Similarity 70.1%; Pred. No. 8.9e-95;  
Matches 329; Conservative 62; Mismatches 68; Indels 10; Gaps 4;

```

OY 1 MSHRAARRKASVTDLYKTCCKOSGCPSPDVNKEGTTLADKTLQWSSLGIFLGIG 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MSHRAARRKASVTDLYKTCCKOSGCPSPDVNKEGTTLADKTLQWSSLGIFLGIG 60

OY 61 TSGGTGRTGYIPLGGSNTVDVGPTRPVVIEPVGPTDPSITVLTIEDSSVVTSGARP 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 TSGGTGRTGYIPLGGSNTVDVGPTRPVVIEPVGPTDPSITVLTIEDSSVVTSGARP 120

OY 121 TFGTSGFDITSAGTTTTPAVLDITPSSTSVISTNTNTNPAFSPSIIIEVPQTEVSGNV 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 TFGTSGFEITSSSTTTPAVLDITPSSTSVISTNTNTNPAFSPSIIIEVPQTEVSGNV 180

OY 181 FVGTPGTGTYEIEIPQTRASSSGTEPISSPTPLPVRRVAGRLXSRAGQVRSVANPE 240
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 181 FVGTPGTGTYEIEIPQTRASSSGTEPISSPTPLPVRRVAGRLXSRAGQVRSVNF 240

OY 241 FLTRPSSLITYDNPAPFVDTTLTFEPRSNVPSDFMDIIRLHRPALTSRRGTVRSRLG 300
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 241 FLTRPSSLITYDNPAPFVDTTLTFEPRSNVPSDFMDIIRLHRPALTSRRGTVRSRV 300

OY 301 ORATMFRSGTQIGARVHFHDISPIASPEYIELQPLVSTENGFLDYADIADIDP 357
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 301 ORATMFRSGTQIGARVHFHDISPIASPEYIELQPLVSTENGFLDYADIADIDP 359

OY 358 AMPV-----SRPTSSAVSYSPYSS--ASSYSNVVPLTSSMDVPVYTGPDITLPTSV 412
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 360 TTVLDIAFHNATFTRSHISVPSLASSTYATNTPIGTAMTPTNTGPDVLTPTSP 419

OY 413 WPIYSPAPASTQY-IGIHGTHYLLPPLYEIPKRRRVYFFADGFVA 460
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 420 QLPFTPTPTDITTAITTYGNTYLLPLFLKRRRLYFFADGIVA 468

```

## RESULT 4

P2ML39  
L2 protein - human papillomavirus type 39  
C:Species: human papillomavirus type 39  
A:Note: host Homo sapiens (man)  
C>Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 16-Jul-1999  
C:Accession: G38502  
R:Volpers, C.; Strebeck, R.E.  
Virology 181, 419-423, 1991  
A:Title: Genome organization and nucleotide sequence of human papillomavirus type 39.  
A:Reference number: A38502; MUID:91135017; PMID:1847266  
A:Accession: G38502  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-470 <VOL>  
A:Cross-references: GB:M62849; EMBL:M38185; NID:9333245; PIDN:AAA47055.1; PID:g463191  
C:Superfamily: papillomavirus L2 protein  
C:Keywords: late protein

Query Match 70.1%; Score 1686.5; DB 1; Length 470;  
Best Local Similarity 69.6%; Pred. No. 5e-93;  
Matches 328; Conservative 60; Mismatches 70; Indels 13; Gaps 7;

```

OY 1 MSHRAARRKASVTDLYKTCCKOSGCPSPDVNKEGTTLADKTLQWSSLGIFLGIG 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 1 MSHRAARRKASVTDLYKTCCKOSGCPSPDVNKEGTTLADKTLQWSSLGIFLGIG 60

OY 61 TSGGTGRTGYIPLGGSNTVDVGPTRPVVIEPVGPTDPSITVLTIEDSSVVTSGARP 120
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 TGTGTGRTGYIPLGGSNTVDVGPTRPVVIEPVGPTDPSITVLTIEDSSVVTSGARP 120

OY 121 TFGTSGFDITSAGTTTTPAVLDITPSSTSVISTNTNTNPAFSPSIIIEVPQTEVSGNV 180
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 121 TFGTSGFEITSSSTTTPAVLDITPSSTSVISTNTNTNPAFSPSIIIEVPQTEVSGNI 180

```











**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 17, 2003, 10:57:04 ; Search time 20.4783 Seconds  
(without alignments)  
952.486 Million cell updates/sec

Title: US-08-913-644-4

Perfect score: 2405  
Sequence: 1 MWSHRARRRRKASVTDLTKT.....FIPKRRKRVPEFADGFVA 461

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 1008  
Listing first 45 summaries

Database: Issued Patents, AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PCRTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/Backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2405	100.0	461	2	US-08-409-122-4
2	2405	100.0	461	2	US-08-408-669-4
3	2317.5	96.4	462	4	US-09-520-822A-4
4	1134	47.2	601	2	US-08-606-288-7
5	1134	47.2	601	2	US-08-606-288-10
6	1134	47.2	601	3	US-09-347-483-7
7	1134	47.2	601	3	US-09-347-483-10
8	1133.5	47.1	473	4	US-09-520-822A-2
9	1130.5	47.0	473	2	US-08-700-670A-23
10	408.5	17.0	469	4	US-09-479-645A-6
11	408.5	17.0	469	4	US-09-479-645A-8
12	380.5	15.8	964	4	US-08-484-791-2
13	151	6.3	978	4	US-09-556-706B-2
14	151	6.3	907	3	US-08-783-774-2
15	151	6.3	907	4	US-09-328-599A-1
16	151	6.3	907	5	PCT-US95-04611A-19
17	148	6.2	2972	4	US-09-579-181-2
18	148	6.2	3118	4	US-09-579-181-1
19	145	6.0	2736	4	US-09-252-991A-30227
20	144.5	6.0	1537	1	US-08-325-267A-2
21	133	5.5	669	4	US-09-071-035-264
22	133	5.5	1638	4	US-09-071-035-258
23	133	5.5	1638	4	US-09-071-035-262
24	133	5.5	1638	4	US-09-071-035-266
25	131.5	5.5	941	4	US-07-757-022B-14
26	131.5	5.5	1022	4	US-07-757-022B-84
27	131.5	5.5	1038	4	US-07-757-022B-74

28	131.5	5.5	1049	4	US-07-757-022B-58	Sequence 58, Appl
29	131.5	5.5	1140	4	US-07-757-022B-104	Sequence 104, App
30	131.5	5.5	1270	4	US-07-757-022B-44	Sequence 44, Appl
31	131.5	5.5	1311	4	US-07-757-022B-142	Sequence 142, Appl
32	131.5	5.5	1313	4	US-07-757-022B-50	Sequence 50, Appl
33	131.5	5.5	1314	4	US-07-757-022B-46	Sequence 46, Appl
34	131.5	5.5	1320	4	US-07-757-022B-60	Sequence 60, Appl
35	131.5	5.5	1354	4	US-07-757-022B-48	Sequence 48, Appl
36	131.5	5.5	1361	4	US-07-757-022B-52	Sequence 52, Appl
37	131.5	5.5	1404	4	US-07-757-022B-62	Sequence 62, Appl
38	131.5	5.5	1404	4	US-09-328-352-5503	Sequence 5503, App
39	130	5.4	1132	4	US-09-198-452A-466	Sequence 466, App
40	127	5.3	1447	3	PCT-US94-05277-2	Sequence 25, Appl
41	126.5	5.3	1447	5	PCT-US94-05277-2	Sequence 2, Appl
42	126.5	5.3	1601	4	US-09-345-473E-40	Sequence 40, Appl
43	126.5	5.3				
44	126.5	5.3				
45	126.5	5.3				

#### ALIGNMENTS

RESULT 1  
US-08-409-122-4  
Sequence 4, Application US/08409122  
Patent No. 5820870  
GENERAL INFORMATION:  
APPLICANT: JOYCE, JAMES G.  
APPLICANT: GEORGE, HUGH A.  
APPLICANT: HOFMANN, KATHRYN J.  
APPLICANT: JANSSEN, KATHRYN U.  
APPLICANT: NEPPER, MICHAEL P.  
TITLE OF INVENTION: DNA ENCODING HUMAN PAPILLOMAVIRUS TYPE 18 VACCINE  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CHRISTINE E. CARTY - MERCK & CO., INC.  
STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000  
CITY: RAHWAY  
STATE: NJ  
COUNTRY: US  
ZIP: 07065-0907  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSRO Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/409,122  
CLASSIFICATION: 435  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/408,669  
FILING DATE: 22-MAR-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: CARTY, CHRISTINE E  
REGISTRATION NUMBER: 36,099  
REFERENCE/DOCKET NUMBER: 19425  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 908-594-6734  
TELEFAX: 908-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 461 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:

US-08-409-122-4

Query Match 100.0%; Score 2405; DB 2; Length 461;  
Best Local Similarity 100.0%; Pred. No. 7.3e-194;  
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHRAARRKRASTYDLYKTKQSGTSPDVNKNVEGTTLADKILQWSSLGIFLGIG 60  
DB 1 MSHRAARRKRASTYDLYKTKQSGTSPDVNKNVEGTTLADKILQWSSLGIFLGIG 60  
QY 61 TSGGTGRTGYIPLGGRSNTVVDGPTRPVIEPVGPTDPSITLLIEDSSVTVSGAPR 120  
DB 61 TSGGTGRTGYIPLGGRSNTVVDGPTRPVIEPVGPTDPSITLLIEDSSVTVSGAPR 120  
QY 121 TFGTSGFDITSAGTTTPAVLDITPSSTSVSISTNFTNPAFSDPSIIEVQGEVSGNV 180  
DB 121 TFGTSGFDITSAGTTTPAVLDITPSSTSVSISTNFTNPAFSDPSIIEVQGEVSGNV 180  
QY 181 FVGPTSGTHGEIEIPLQTFASSGTGEPISTPLPVRRVAGRLYSRAVQVSVANPE 240  
DB 181 FVGPTSGTHGEIEIPLQTFASSGTGEPISTPLPVRRVAGRLYSRAVQVSVANPE 240  
QY 241 FLTRPSSLIYYDNAPFEPVDTTLTFEPRSNDPDSDFMDIIRLHRPALTSRGTFRSRLG 300  
DB 241 FLTRPSSLIYYDNAPFEPVDTTLTFEPRSNDPDSDFMDIIRLHRPALTSRGTFRSRLG 300  
QY 301 QRATMFTRSGTQICARHFYHDISPAPSPYIELQPLVSATENGLEFDIYADDIDPAMP 360  
DB 301 QRATMFTRSGTQICARHFYHDISPAPSPYIELQPLVSATENGLEFDIYADDIDPAMP 360  
QY 361 VPSRPTTSSAVSTYSPITSSASSYSNTVPLTSSMDVPVYTGPDITLPPTSVMPVIVSPTA 420  
DB 361 VPSRPTTSSAVSTYSPITSSASSYSNTVPLTSSMDVPVYTGPDITLPPTSVMPVIVSPTA 420  
QY 421 PASTQYIGIGHGTHYLLMPLYYFIPKRRKRVYFFADGFVAA 461  
DB 421 PASTQYIGIGHGTHYLLMPLYYFIPKRRKRVYFFADGFVAA 461

RESULT 2

US-08-408-669-4  
Sequence 4, Application US/08408669  
Patent No. 5840306  
GENERAL INFORMATION:  
APPLICANT: HOFMANN, KATHRYN J.  
APPLICANT: JANSEN, KATHRYN U.  
APPLICANT: NEEBER, MICHAEL P.  
TITLE OF INVENTION: DNA ENCODING HUMAN PAPILLOMAVIRUS TYPE 18  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CHRISTINE E. CARTY - MERCK & CO., INC.  
STREET: 126 EAST LINCOLN AVENUE - P.O. BOX 2000  
CITY: RAHWAY  
STATE: NJ  
COUNTRY: US  
ZIP: 07065-0907  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 1.5  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/408,669  
FILING DATE: 22-MAR-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: CARTY, CHRISTINE E  
REGISTRATION NUMBER: 36,099  
REFERENCE/DOCKET NUMBER: 19424  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 908-594-6734  
TELEFAX: 908-594-4720  
TELEX:  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 461 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
US-08-408-669-4

Query Match 100.0%; Score 2405; DB 2; Length 461;  
Best Local Similarity 100.0%; Pred. No. 7.3e-194;  
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MSHRAARRKRASTYDLYKTKQSGTSPDVNKNVEGTTLADKILQWSSLGIFLGIG 60  
DB 1 MSHRAARRKRASTYDLYKTKQSGTSPDVNKNVEGTTLADKILQWSSLGIFLGIG 60  
QY 61 TSGGTGRTGYIPLGGRSNTVVDGPTRPVIEPVGPTDPSITLLIEDSSVTVSGAPR 120  
DB 61 TSGGTGRTGYIPLGGRSNTVVDGPTRPVIEPVGPTDPSITLLIEDSSVTVSGAPR 120  
QY 121 TFGTSGFDITSAGTTTPAVLDITPSSTSVSISTNFTNPAFSDPSIIEVQGEVSGNV 180  
DB 121 TFGTSGFDITSAGTTTPAVLDITPSSTSVSISTNFTNPAFSDPSIIEVQGEVSGNV 180  
QY 181 FVGPTSGTHGEIEIPLQTFASSGTGEPISTPLPVRRVAGRLYSRAVQVSVANPE 240  
DB 181 FVGPTSGTHGEIEIPLQTFASSGTGEPISTPLPVRRVAGRLYSRAVQVSVANPE 240  
QY 241 FLTRPSSLIYYDNAPFEPVDTTLTFEPRSNDPDSDFMDIIRLHRPALTSRGTFRSRLG 300  
DB 241 FLTRPSSLIYYDNAPFEPVDTTLTFEPRSNDPDSDFMDIIRLHRPALTSRGTFRSRLG 300  
QY 301 QRATMFTRSGTQICARHFYHDISPAPSPYIELQPLVSATENGLEFDIYADDIDPAMP 360  
DB 301 QRATMFTRSGTQICARHFYHDISPAPSPYIELQPLVSATENGLEFDIYADDIDPAMP 360  
QY 361 VPSRPTTSSAVSTYSPITSSASSYSNTVPLTSSMDVPVYTGPDITLPPTSVMPVIVSPTA 420  
DB 361 VPSRPTTSSAVSTYSPITSSASSYSNTVPLTSSMDVPVYTGPDITLPPTSVMPVIVSPTA 420  
QY 421 PASTQYIGIGHGTHYLLMPLYYFIPKRRKRVYFFADGFVAA 461  
DB 421 PASTQYIGIGHGTHYLLMPLYYFIPKRRKRVYFFADGFVAA 461

RESULT 3

US-09-520-822A-4  
Sequence 4, Application US/09520822A  
Patent No. 6551597  
GENERAL INFORMATION:  
APPLICANT: Harrison, Stephen  
APPLICANT: Chen, Xiaojiang  
TITLE OF INVENTION: Vaccine compositions and methods for human papilloma  
virus  
FILE REFERENCE: Hayard/Harrison 12687/1120  
CURRENT APPLICATION NUMBER: US/09/520,822A  
CURRENT FILING DATE: 2000-03-08  
PRIOR APPLICATION NUMBER: 60/125208  
PRIOR FILING DATE: 1999-03-18  
PRIOR APPLICATION NUMBER: 60/148544  
PRIOR FILING DATE: 1999-08-12  
NUMBER OF SEQ ID NOS: 4  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4  
LENGTH: 462

TYPE: PR1  
ORGANISM: Human papillomavirus type 18  
US-08-520-822A-4

Query Match 96.4%; Score 2317.5; DB 4; Length 462;  
Best Local Similarity 96.8%; Pred. No. 1.6e-186;  
Matches 447; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 1 MYSHRAARRKASVVDLYKTKQSGTCDPVDVYKVEGTLADKILQWSSLGIFLGSLGIG 60  
DB 1 MYSHRAARRKASVVDLYKTKQSGTCDPVDVYKVEGTLADKILQWSSLGIFLGSLGIG 60  
QY 61 TSSGSGRTGYIPLGSRNTVVDGTPRPVIEVGPDPSPILVLIEDSSVVTSGAPRP 120  
DB 61 TSSGSGRTGYIPLGSRNTVVDGTPRPVIEVGPDPSPILVLIEDSSVVTSGAPRP 120  
QY 121 TPTGSGPDITAGTTPAVLDITPSTSVSISTNFNPAPSDPSIIEVPTGEGVAGN 180  
DB 121 TPTGSGPDITAGTTPAVLDITPSTSVSISTNFNPAPSDPSIIEVPTGEGVAGN 180  
QY 181 FVGTPSTGTHGEIEIPLQTFASSGTGEPISSSTPLPTVRRVAGPRLYSRAYOVSANPE 240  
DB 181 FVGTPSTGTHGEIEIPLQTFASSGTGEPISSSTPLPTVRRVAGPRLYSRAYOVSANPE 240  
QY 241 FLTRPSSLITYNPAPEVDITLTPEPRSNVDPDSDFMDIIRLRPALTSRGTVRSRLG 300  
DB 241 FLTRPSSLITYNPAPEVDITLTPEPRSNVDPDSDFMDIIRLRPALTSRGTVRSRLG 300  
QY 301 ORATPSTRSGQIGARVHFYHDISPAPSEYIEIOLPLVSATEDNGLFTIYADDIDPAMP 360  
DB 301 ORATPSTRSGQIGARVHFYHDISPAPSEYIEIOLPLVSATEDNGLFTIYADDIDPAMP 360  
QY 361 VPSRPTSSAVSTPSSISSASSYSNVYPLTSSMDVPYTGPDITLP-PTSAWPIVSP 419  
DB 361 VPSRPTSSAVSTPSSISSASSYSNVYPLTSSMDVPYTGPDITLP-PTSAWPIVSP 419  
QY 420 APASTOYIGIHGTHYLMPLFYIPKRRKRVYFFADGVAA 461  
DB 420 APASTOYIGIHGTHYLMPLFYIPKRRKRVYFFADGVAA 461

## RESULT 4

US-08-606-288-7  
Sequence 7, Application US/08606288

GENERAL INFORMATION:  
PATENT NO. 5955087  
APPLICANT: Whittle, N.R.  
APPLICANT: Carmichael, J.P.  
APPLICANT: Connor, S.E.  
APPLICANT: Thompson, H.S.G.  
TITLE OF INVENTION: Polypeptides Useful as Immunotherapeutic  
TITLE OF INVENTION: Agents, and Methods of Polypeptide Preparation  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Suite 3400, Four Embarcadero Center  
CITY: San Francisco  
STATE: California  
COUNTRY: USA

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
FILING DATE: 23-FEB-1996  
APPLICATION NUMBER: US/08/606,288  
PRIORITY APPLICATION DATA:  
FILING DATE: 24-FEB-1995  
APPLICATION NUMBER: GB 9503786.7  
PRIORITY APPLICATION DATA:  
FILING DATE: 24-FEB-1995  
APPLICATION NUMBER: US 60/000034

FILING DATE: 08-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9515478.7  
FILING DATE: 28-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Walter H. Dreger  
REGISTRATION NUMBER: 24,190  
REFERENCE/DOCKET NUMBER: A-63284/MHD  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 601 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-606-288-7

Query Match 47.2%; Score 1134; DB 2; Length 601;  
Best Local Similarity 50.8%; Pred. No. 6.2e-87;  
Matches 236; Conservative 75; Mismatches 136; Indels 18; Gaps 10;

QY 5 RAARRRASVTDLYKTKQSGTCDPVDVYKVEGTLADKILQWSSLGIFLGSLGIGTSG 64  
DB 35 RAARRRASVTDLYKTKQSGTCDPVDVYKVEGTLADKILQWSSLGIFLGSLGIGTSG 64  
QY 65 TGRGTGYIPLGSRNTVVDGTPRPVIEVGPDPSPILVLIEDSSVVTSGAPRPPT 123  
DB 95 TGRGTGYIPLGSRNTVVDGTPRPVIEVGPDPSPILVLIEDSSVVTSGAPRPPT 123  
QY 124 GTSGPDITAGTTPAVLDITPSTSVSISTNFNPAPSDPSIIEVPTGEGVAGN 183  
DB 155 AHGFTITSETTTPAILDVSVTSHT---TSTIFRPVTEBSVYOPQPPVANGHILIS 211  
QY 184 TPTSGTHGEIEIPLQTFASSGTGEPISSSTPLPTVRRVAGPRLYSRAYOVSANPE 240  
DB 212 APTTSHPIEIEIPLDFVLISSSDGPTSTPVGTA---PRVGLYSRAHQVYTDPA 268  
QY 241 FLTRPSSLITYNPAPEVDITLTPEPRSNVDPDSDFMDIIRLRPALTSRGTVRSRLG 298  
DB 269 FLSTPQRLITYNPAPEVDITLTPEPRSNVDPDSDFMDIIRLRPALTSRGTVRSRLG 328  
QY 299 LGORATPSTRSGQIGARVHFYHDISPAPSEYIEIOLPLVSATEDNGLFTIYADDIDPA 358  
DB 329 IGORSMHTRSGHIGARHIFDYDISPIQAAEIEMHVLVAQDET---FDIYASFED 386  
QY 359 MEVPSRPTSSAVSTPSSISSASSYSNVYPLTSSMDVPYTGPDITLP-PTSAWPIVSP 417  
DB 387 INPTQHPVYNISDTYLTSTPNVTQPMGNTVPLSIPNDLFLQSGPDITFPAPMGTPPS 446  
QY 418 PTAPA-STOYIGIHGTHYLMPLFYIPKRRKRVYFFADGVAA 461  
DB 447 PVTPLPTGPVFTSGGYLHPAWYFARRRRRIPLFSD--VAA 489

## RESULT 5

US-08-606-288-10  
Sequence 10, Application US/08606288

GENERAL INFORMATION:  
PATENT NO. 5955087  
APPLICANT: Whittle, N.R.  
APPLICANT: Carmichael, J.P.  
APPLICANT: Connor, S.E.  
APPLICANT: Thompson, H.S.G.  
TITLE OF INVENTION: Polypeptides Useful as Immunotherapeutic  
TITLE OF INVENTION: Agents, and Methods of Polypeptide Preparation  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Suite 3400, Four Embarcadero Center  
CITY: San Francisco

STATE: California  
COUNTRY: USA  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/606,288  
FILING DATE: 23-FEB-1996  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9503786.7  
FILING DATE: 24-FEB-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/000034  
FILING DATE: 08-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9515478.7  
FILING DATE: 28-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Walter H. Dreger  
REGISTRATION NUMBER: 24,190  
REFERENCE/DOCKET NUMBER: A-63284/MHD  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 601 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
US-08-606-288-10

Query Match 47.28; Score 1134; DB 2; Length 601;  
Best Local Similarity 50.88; Pred. No. 6.2e-87;  
Matches 236; Conservative 75; Mismatches 136; Indels 18; Gaps 10;

QY 5 RAARRKRAASYTDLYKTKOSGTCPSDVYKNEGTTLADKILQWSSLCIFLGAGIGTSGS 64  
DB 35 RAARRKRAASYTDLYKTKOSGTCPSDVYKNEGTTLADKILQWSSLCIFLGAGIGTSGS 94  
QY 65 TGGRTGYIPLGKSNVVDVGP--TRPVVIEPVGPDPSTVTLIEDSSVVTSGAPRPTFT 123  
DB 95 TGGRTGYIPLGKSNVVDVGP--TRPVVIEPVGPDPSTVTLIEDSSVVTSGAPRPTFT 154  
QY 124 GTSGFDITSAAGTTTTPAVLDITPSTSVSISTNTFTNPAFSDPSLIEVPQGEVSGANFVG 183  
DB 155 AHGFTITSETTTPALDVSYSHT---TTSIFRNVFTFEPSTVQOPPEVANGHILIS 211  
QY 184 TPTSGHGYEIPLQTFPASSGTEGPSTPLP--TVRRVAGPR--LYSRAYQGVSVANPE 240  
DB 212 APTTSHPIEIPIDTDFVSISSDSGPTSPVPCTABR---PRVGLYSRALHQVYVDPDA 268  
QY 241 FLTRPSSLITYDNPAPFVDTITLFEPRS--NVPDSFMDIIRLRALTSRGTVAFRSR 298  
DB 269 FLTRPSSLITYDNPAPFVDTITLFEPRS--NVPDSFMDIIRLRALTSRGTVAFRSR 328  
QY 299 LGRATMFTSSGTOIGARVHFYHDIISPIAPSEYIELQPLVSATEDGLFDIYADDIDPA 358  
DB 329 IGRGSHMTRSGKHIGKIRHIFYDISPIAQAEEIEHMLPVAAOEDT--EDITAKSTEPD 386  
QY 359 MPVPSRPTSSAVSVSPITSSASS--YSNVTVPLTSSMDVPVYTGPDITLTPPTSVMPFVS 417  
DB 387 INPQHVVTNISTDYLSTFNTVQOPGNGNTVPLSINDLEFGSDPITPTPLANMGTPFS 446  
QY 418 PTAPA-STOYIGHGHYIYLMPLYYIPKKRRKRVYFADGFVAA 461  
DB 447 PVTPALPTGPVITGSGFYLHPAMYFARKRRKRRIPLFFSD--VAA 489

RESULT 6  
US-09-347-483-7  
Sequence 7, Application US/09347483  
Patent No. 6123948  
GENERAL INFORMATION:  
APPLICANT: Whittle, N.R.  
APPLICANT: Carmichael, J.P.  
APPLICANT: Connor, S.E.  
APPLICANT: Thompson, H.S.G.  
APPLICANT: Wilson, M.J.  
TITLE OF INVENTION: Polypeptides Useful as Immunotherapeutic Agents, and Methods of Polypeptide Preparation  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Suite 3400, Four Embarcadero Center  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/347,483  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/606,288  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/000034  
FILING DATE: 08-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9515478.7  
FILING DATE: 28-JUL-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Walter H. Dreger  
REGISTRATION NUMBER: 24,190  
REFERENCE/DOCKET NUMBER: A-63284/MHD  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 601 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-347-483-7

Query Match 47.28; Score 1134; DB 3; Length 601;  
Best Local Similarity 50.88; Pred. No. 6.2e-87;  
Matches 236; Conservative 75; Mismatches 136; Indels 18; Gaps 10;

QY 5 RAARRKRAASYTDLYKTKOSGTCPSDVYKNEGTTLADKILQWSSLCIFLGAGIGTSGS 64  
DB 35 RAARRKRAASYTDLYKTKOSGTCPSDVYKNEGTTLADKILQWSSLCIFLGAGIGTSGS 94  
QY 65 TGGRTGYIPLGKSNVVDVGP--TRPVVIEPVGPDPSTVTLIEDSSVVTSGAPRPTFT 123  
DB 95 TGGRTGYIPLGKSNVVDVGP--TRPVVIEPVGPDPSTVTLIEDSSVVTSGAPRPTFT 154  
QY 124 GTSGFDITSAAGTTTTPAVLDITPSTSVSISTNTFTNPAFSDPSLIEVPQGEVSGANFVG 183  
DB 155 AHGFTITSETTTPALDVSYSHT---TTSIFRNVFTFEPSTVQOPPEVANGHILIS 211  
QY 184 TPTSGHGYEIPLQTFPASSGTEGPSTPLP--TVRRVAGPR--LYSRAYQGVSVANPE 240  
DB 212 APTTSHPIEIPIDTDFVSISSDSGPTSPVPCTABR---PRVGLYSRALHQVYVDPDA 268  
QY 241 FLTRPSSLITYDNPAPFVDTITLFEPRS--NVPDSFMDIIRLRALTSRGTVAFRSR 298

Db	269	FLSTFQRLITLYDNPNYIEGGDVSVOGSHSINHNADEAFMDIIRLHRPALASRGLYKSR	3228
Qy	299	LGQRATMTFRSTQIGARVHYEHDISPIAPSPSEYIELOPLVSATEDNGLEDIYADDIPA	358
Db	329	IGORSMTSRGSKHIGARHYEHYDIPISQAAMEIEHMLPVAQAEDT--FDIYAKSPEPD	366
Qy	359	MPVSRPPTSSAVSTYSPTISSASS--YSNVYVPLPSSMWVPLYTGDIILPLPSSWPIYS	417
Db	387	INPQHPATNINSDIYLTSTPNTVTPQWGTATYPLSPINDLFLQSGDPTTFPTAPMGTPFS	446
Qy	418	PTAFA-STQYIGIHGTHYLMPLYEIPKKRRKRVPEFGDGVAA	461
Db	447	PVTPLALPTGPVFTTCSGFTYLPAMVFAKKRRKRIRLPEESD--VAA	489

```

1      RESULT 7
2      US-09-347-483-10
3      : Sequence 10, Application US/09347483
4      : Patent No. 6123948
5      : GENERAL INFORMATION:
6      : APPLICANT: Whittle, N.R.
7      : APPLICANT: Carmichael, J.P.
8      : APPLICANT: Connor, S.E.
9      : APPLICANT: Thompson, H.S.G.
10     : APPLICANT: Wilson, M.J.
11     : TITLE OF INVENTION: Polypeptides Useful as Immunotherapeutic
12     : TITLE OF INVENTION: Agents, and Methods of Polypeptide Preparation
13     : NUMBER OF SEQUENCES: 10
14     : CORRESPONDENCE ADDRESS:
15     : ADDRESS: Flehr, Hobbach, Test, Albritton & Herbert
16     : STREET: Suite 3400, Four Embarradero Center
17     : CITY: San Francisco
18     : STATE: California
19     : COUNTRY: USA
20     : COMPUTER READABLE FORM:
21     : MEDIUM TYPE: Floppy disk
22     : COMPUTER: IBM PC compatible
23     : OPERATING SYSTEM: PC-DOS/MS-DOS
24     : SOFTWARE: Patentln Release #1.0, Version #1.25 (EPO)
25     : CURRENT APPLICATION DATA:
26     : APPLICATION NUMBER: US/09/347,483
27     : FILING DATE:
28     : PRIOR APPLICATION DATA:
29     : APPLICATION NUMBER: US 08/606,288
30     : FILING DATE:
31     : PRIOR APPLICATION DATA:
32     : APPLICATION NUMBER: US 60/000034
33     : FILING DATE: 08-JUN-1995
34     : PRIOR APPLICATION DATA:
35     : APPLICATION NUMBER: GB 9515478.7
36     : FILING DATE: 28-JUL-1995
37     : ATTORNEY/AGENT INFORMATION:
38     : NAME: Walter H. Dreyer
39     : REGISTRATION NUMBER: 24,190
40     : REFERENCE/DOCKET NUMBER: A-63284/WHD
41     : TELECOMMUNICATION INFORMATION:
42     : TELEPHONE: (415) 781-1989
43     : TELEFAX: (415) 398-3249
44     : INFORMATION FOR SEQ ID NO: 10:
45     : SEQUENCE CHARACTERISTICS:
46     : LENGTH: 601 amino acids
47     : TYPE: amino acid
48     : TOPOLOGY: linear
49     : MOLECULE TYPE: protein
50     : HYPOTHEITICAL: NO
51     : ANTI-SENSE: NO
52     :
53     :
54     :
55     :
56     :
57     :
58     :
59     :
60     :
61     :
62     :
63     :
64     :
65     :
66     :
67     :
68     :
69     :
70     :
71     :
72     :
73     :
74     :
75     :
76     :
77     :
78     :
79     :
80     :
81     :
82     :
83     :
84     :
85     :
86     :
87     :
88     :
89     :
90     :
91     :
92     :
93     :
94     :
95     :
96     :
97     :
98     :
99     :
100    :
101    :
102    :
103    :
104    :
105    :
106    :
107    :
108    :
109    :
110    :
111    :
112    :
113    :
114    :
115    :
116    :
117    :
118    :
119    :
120    :
121    :
122    :
123    :
124    :
125    :
126    :
127    :
128    :
129    :
130    :
131    :
132    :
133    :
134    :
135    :
136    :
137    :
138    :
139    :
140    :
141    :
142    :
143    :
144    :
145    :
146    :
147    :
148    :
149    :
150    :
151    :
152    :
153    :
154    :
155    :
156    :
157    :
158    :
159    :
160    :
161    :
162    :
163    :
164    :
165    :
166    :
167    :
168    :
169    :
170    :
171    :
172    :
173    :
174    :
175    :
176    :
177    :
178    :
179    :
180    :
181    :
182    :
183    :
184    :
185    :
186    :
187    :
188    :
189    :
190    :
191    :
192    :
193    :
194    :
195    :
196    :
197    :
198    :
199    :
200    :
201    :
202    :
203    :
204    :
205    :
206    :
207    :
208    :
209    :
210    :
211    :
212    :
213    :
214    :
215    :
216    :
217    :
218    :
219    :
220    :
221    :
222    :
223    :
224    :
225    :
226    :
227    :
228    :
229    :
230    :
231    :
232    :
233    :
234    :
235    :
236    :
237    :
238    :
239    :
240    :
241    :
242    :
243    :
244    :
245    :
246    :
247    :
248    :
249    :
250    :
251    :
252    :
253    :
254    :
255    :
256    :
257    :
258    :
259    :
260    :
261    :
262    :
263    :
264    :
265    :
266    :
267    :
268    :
269    :
270    :
271    :
272    :
273    :
274    :
275    :
276    :
277    :
278    :
279    :
280    :
281    :
282    :
283    :
284    :
285    :
286    :
287    :
288    :
289    :
290    :
291    :
292    :
293    :
294    :
295    :
296    :
297    :
298    :
299    :
300    :
301    :
302    :
303    :
304    :
305    :
306    :
307    :
308    :
309    :
310    :
311    :
312    :
313    :
314    :
315    :
316    :
317    :
318    :
319    :
320    :
321    :
322    :
323    :
324    :
325    :
326    :
327    :
328    :
329    :
330    :
331    :
332    :
333    :
334    :
335    :
336    :
337    :
338    :
339    :
340    :
341    :
342    :
343    :
344    :
345    :
346    :
347    :
348    :
349    :
350    :
351    :
352    :
353    :
354    :
355    :
356    :
357    :
358    :
359    :
360    :
361    :
362    :
363    :
364    :
365    :
366    :
367    :
368    :
369    :
370    :
371    :
372    :
373    :
374    :
375    :
376    :
377    :
378    :
379    :
380    :
381    :
382    :
383    :
384    :
385    :
386    :
387    :
388    :
389    :
390    :
391    :
392    :
393    :
394    :
395    :
396    :
397    :
398    :
399    :
400    :
401    :
402    :
403    :
404    :
405    :
406    :
407    :
408    :
409    :
410    :
411    :
412    :
413    :
414    :
415    :
416    :
417    :
418    :
419    :
420    :
421    :
422    :
423    :
424    :
425    :
426    :
427    :
428    :
429    :
430    :
431    :
432    :
433    :
434    :
435    :
436    :
437    :
438    :
439    :
440    :
441    :
442    :
443    :
444    :
445    :
446    :
447    :
448    :
449    :
450    :
451    :
452    :
453    :
454    :
455    :
456    :
457    :
458    :
459    :
460    :
461    :
462    :
463    :
464    :
465    :
466    :
467    :
468    :
469    :
470    :
471    :
472    :
473    :
474    :
475    :
476    :
477    :
478    :
479    :
480    :
481    :
482    :
483    :
484    :
485    :
486    :
487    :
488    :
489    :
490    :
491    :
492    :
493    :
494    :
495    :
496    :
497    :
498    :
499    :
500    :
501    :
502    :
503    :
504    :
505    :
506    :
507    :
508    :
509    :
510    :
511    :
512    :
513    :
514    :
515    :
516    :
517    :
518    :
519    :
520    :
521    :
522    :
523    :
524    :
525    :
526    :
527    :
528    :
529    :
530    :
531    :
532    :
533    :
534    :
535    :
536    :
537    :
538    :
539    :
540    :
541    :
542    :
543    :
544    :
545    :
546    :
547    :
548    :
549    :
550    :
551    :
552    :
553    :
554    :
555    :
556    :
557    :
558    :
559    :
560    :
561    :
562    :
563    :
564    :
565    :
566    :
567    :
568    :
569    :
570    :
571    :
572    :
573    :
574    :
575    :
576    :
577    :
578    :
579    :
580    :
581    :
582    :
583    :
584    :
585    :
586    :
587    :
588    :
589    :
590    :
591    :
592    :
593    :
594    :
595    :
596    :
597    :
598    :
599    :
600    :
601    :
602    :
603    :
604    :
605    :
606    :
607    :
608    :
609    :
610    :
611    :
612    :
613    :
614    :
615    :
616    :
617    :
618    :
619    :
620    :
621    :

```

[illegible]

```

RESULT 8
US-09-520-822A-2
US-09-520-822A-2
Sequence 2, Application US/09520822A
Patent No. 6551597
GENERAL INFORMATION:
APPLICANT: Harrison, Stephen
APPLICANT: Chen, Xiaojiang
TITLE OF INVENTION: Vaccine compositions and methods for human papilloma
TITLE OF INVENTION: virus
FILE REFERENCE: Harvard/Harrison 12687/1120
CURRENT FILING DATE: US/09/520,822A
CURRENT FILING DATE: 2000-03-08
PRIOR APPLICATION NUMBER: 60/125208
PRIOR FILING DATE: 1999-03-18
PRIOR APPLICATION NUMBER: 60/148544
PRIOR FILING DATE: 1999-08-12
NUMBER OF SEQ ID NOS: 4
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 473
TYPE: PRT
ORGANISM: Human papillomavirus type 16
US-09-520-822A-2

Query Match 47.1%; Score 1133.5; DB 4; Length 473;
Best Local Similarity 49.2%; Pred. No. 4,8e-87;
Matches 237; Conservative 73; Mismatches 131; Indels 41; Gaps 12

OY 6 AARRRASVYDLYTCTCKSGTCPSDVYKVBGTTLADKLTLQWSSLGIFLGIGTGSQT 65
Db 7 AKRRKRASATQTLTKCKAGCTCPPLIKVEGKRTIAEQILQYSGMVGFEFGIGTGSQT 66
OY 66 GGGTGYIPLGGRSTMYVD-VGPTRPVYIEVPGPRDSIVLILIEDSSVYTSAPR--PTE 122
Db 67 GGRGTGYIPLGRRPPLADYTLAPVRPRPLTVDPVGPSPISVLSVEETSTFIDAGAPTSVPSI 126
OY 123 -TGTSGFDITSAGTTTPAVLDITFSSTSISTFTFTPAFSDPSIIIEVPGTGEVSGNVE 181
Db 127 PPDVSGSITSTSTPTPAIDLDINMTVTV-----TTHNNPTTDSVYLPPTPAETGCHFT 182
OY 182 VGPTSGTGHGVEEILPLQTFASSGCGEEDISSPTLPTRVRVAGPRLYSRAYQOVSANPEF 241

```

Db 183 LSSSTISHNHEELIPMTFIYSTNPNTVTSSTPIPGSRPARLGLYSTTQOVVDPAF 242  
QY 242 LTRPSLITYDNPAFE--PVDTLTFEPRSN---VPSDPMIIRLRPALTSRGTVR 295  
Db 243 VTPPTKLITYDNPAFEGIDVNTLYFSSNDNSINIADDPDLIVALHRLPALTSRRGIR 302  
QY 296 FSRIGQATMTFRGCTQIGARVHEHYNDISPIAPSEYIELQPLVSAE----- 343  
Db 303 YSRIGNKQTLKTRSGKSGIAKAVHYLDSTIDPAEE-IELOTIPSTYTTTSHAASPTSI 361  
QY 344 DNGLEFDIYADDI--DPAMPVPSRPTSSAVSTSPITSSASSYNTVPLTSSMDVPY 400  
Db 362 NNGLYDIYADDFITDSTTPVPSVPS--LSGIIP-----ANTTIPGCAVNIPLV 411  
QY 401 TGPDTITLPPTSVWDIVSPAPASTQY-IGIHGTHYIYMLPYFIIPKKRRVPEFADGFV 459  
Db 412 SGDPIDPINTDOAPSLPIVPGSPQYTIADAGDFYLPSPYMLRKRKRRLPYEFSOVL 471  
QY 460 AA 461  
Db 472 AA 473

## RESULT 9

US-08-700-670A-23  
Sequence 23, Application US/08700670A

Patent No. 5993821  
GENERAL INFORMATION:  
APPLICANT: FRAZER, Ian  
APPLICANT: ZHOU, Jian  
TITLE OF INVENTION: MODIFIED PAPILLOMA VIRUS L2 PROTEIN AND  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIR: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/700,670A  
FILING DATE: 30-JUL-1996  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/AU95/00043  
FILING DATE: 31-JAN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: AU PM3588  
FILING DATE: 31-JAN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: SAKE, Bernhard D.  
REGISTRATION NUMBER: 28,665  
REFERENCE/DOCKET NUMBER: 65064/107  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 23:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 473 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-700-670A-23

Query Match

47.0%; Score 1130.5; DB 2; Length 473;

Best Local Similarity 49.0%; Pred. No. 8.6e-87;  
Matches 236; Conservative 74; Mismatches 131; Indels 41; Gaps 12;  
QY 6 AARRKASVTLTYKCKSGCTGCPDYNKVGTTLADKILLOWSGIFGLGIGTSGT 65  
Db 7 AKRRKASATLYTCKKAGCTCPDIPKVGKIAEILLYGTMGVFGGLGIGTSGT 66  
QY 66 GGRGTIYPLGGRSNVVD-VGPTRPVYIEVPGTDSIVTLIDSSVYTSAGAR--PTF 122  
Db 67 GGRGTIYPLGGRPTADTLAPVRPPLIVDPVSPDSIVLVEETSPIDAGAVTSPSI 126  
QY 123 -TGSGDITSGAGTTTAVVLDITPSSVSISTNFTNPASDSEIIEVPGTGVSNVF 181  
Db 127 PPDVSGFSITVSTTTPALIDINNTVTV---TTHNPPFTDSVLOPPTAGTGHFT 182  
QY 182 VGTPTSGTHGYEELIPLQTFASSGTGEEPISTPLTVRVARVAPPLRAYQOVANPEF 241  
Db 183 LSSSTISHNHEELIPMTFIYSTNPNTVTSSTPIPGSRPARLGLYSTTQOVVDPAF 242  
QY 242 LTRPSLITYDNPAFE--PVDTLTFEPRSN---VPSDPMIIRLRPALTSRGTVR 295  
Db 243 VTPPTKLITYDNPAFEGIDVNTLYFSSNDNSINIADDPDLIVALHRLPALTSRRGIR 302  
QY 296 FSRIGQATMTFRGCTQIGARVHEHYNDISPIAPSEYIELQPLVSAE----- 343  
Db 303 YSRIGNKQTLKTRSGKSGIAKAVHYLDSTIDPAEE-IELOTIPSTYTTTSHAASPTSI 361  
QY 344 DNGLEFDIYADDI--DPAMPVPSRPTSSAVSTSPITSSASSYNTVPLTSSMDVPY 400  
Db 362 NNGLYDIYADDFITDSTTPVPSVPS--LSGIIP-----ANTTIPGCAVNIPLV 411  
QY 401 TGPDTITLPPTSVWDIVSPAPASTQY-IGIHGTHYIYMLPYFIIPKKRRVPEFADGFV 459  
Db 412 SGDPIDPINTDOAPSLPIVPGSPQYTIADAGDFYLPSPYMLRKRKRRLPYEFSOVL 471  
QY 460 AA 461  
Db 472 AA 473

## RESULT 10

US-09-479-645A-6  
Sequence 6, Application US/09479645A

Patent No. 6489141  
GENERAL INFORMATION:  
APPLICANT: FRAZER, Ian Hector  
APPLICANT: ZHOU, Jian  
TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND METHOD FOR SELECTIVELY  
FILE REFERENCE: 210338.0001/105  
CURRENT APPLICATION NUMBER: US/09/479,645A  
FILING DATE: 2000-01-07  
PRIOR APPLICATION NUMBER: PCT/AU98/00530  
PRIOR FILING DATE: 1998-07-09  
PRIOR APPLICATION NUMBER: AU P07765  
PRIOR FILING DATE: 1997-07-09  
PRIOR APPLICATION NUMBER: AU P09467  
PRIOR FILING DATE: 1997-09-11  
NUMBER OF SEQ ID NOS: 219  
SOFTWARE: patentin Ver. 2.0  
SEQ ID NO 6  
LENGTH: 469  
TYPE: PRT  
ORGANISM: Bovine papillomavirus type 1  
FEATURE:  
OTHER INFORMATION: L2 open reading frame (wild-type)  
US-09-479-645A-6

Query Match 17.0%; Score 408.5; DB 4; Length 469;  
Best Local Similarity 29.3%; Pred. No. 3.8e-26;  
Matches 152; Conservative 59; Mismatches 174; Indels 133; Gaps 24;  
QY 8 RRRKASVTLTYKCKSGCTGCPDYNKVGTTLADKILLOWSGIFGLGIGTGS---- 63

```

Db      6 RKRASAYDLKRTCKQAGCPDPVIRKVEGDTIAKILKFGGLATYLGGLGIGTWSGRV 65
QY      64 GTGRTGYIPL--GGRSNTVVDVGP-----TRPPV-----VIEPVPTPST---VT 105
Db      66 AAGGSPRRYPLRTAGSTSSLASISRAVYAGTRPSIGAGIPLDTLETLGALRPGVEDTV 125
QY      106 LIESSVVTSGAPRPFTGTSGFDITSAGTTP-----AVLDIYP-SST 148
Db      126 LPEAPAIYTPDA----VPADSGLDALSTGDSSTETLITLLEPEGEDIAVLELQPLDRP 181
QY      149 SVSISTNTNPAFSDPSIIEVPQGEVSG--NVFVGPTSGTHGEELIPLQTFASSGCG 206
Db      182 TWQVNAVAHQSSAYHAPLQLO--SIAETSGLENIIEVGGSGLDGDTGGENIELTFGSG---- 236
QY      207 EEPISSTPLPYTRRVAGPRLY--SRAYQOVSVANPEFLRPSLIITYDPAPEPVDTTLT 264
Db      237 --PRISTPRSIASKSGILNMFSKRYTQVPTEDPEVF-----SSQTFANPLIEA----- 284
QY      265 FEPRSNVPSDFMDIIRLHRPALTSRGTVRFSRLGQRATMFTRSQTQIGARVHYHDIS 324
Db      285 -EP-----AVLKGPSGVRGLSQYKKPDLITRSGTEVGPQLHVRYSLS 326
QY      325 PIAPSPETI-----ELQPLVSATEDNGLEDIYADIDPA--MP----- 360
Db      327 TIEDEVALPYVDENTOGIAFVPL--HEEQAGFEIEILDSETHRLLPQNTSSTPGVS 384
QY      361 -----VPSR---PTTSSAVSTY-SPTISSASSYSNVYPLNLSMDVPYTYGDTILPPT 410
Db      365 GVRSLIPLTRESATRPCTGVVYTGSPDYTSAS-----PVTD---POSTSPSLVIDDT 433
QY      411 SVMPIVSPAPASTOYIGIHGTHYVLYLPIYFIPKRRK 448
Db      434 TTTPILII-----IDGHTVDLYSSNYTLHPSLLRKRRK 467

```

```

RESULT 11
US-09-479-645A-8
; Sequence 8, Application US/09479645A
; Patent No. 6489141
; GENERAL INFORMATION:
; APPLICANT: PRAZER, Ian Hector
; APPLICANT: ZHOU, Jian
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCE AND METHOD FOR SELECTIVELY
; TITLE OF INVENTION: EXPRESSING A PROTEIN IN A TARGET CELL OR TISSUE
; FILE REFERENCE: 210338.0001/US
; CURRENT APPLICATION NUMBER: US/09/479,645A
; PRIOR FILING DATE: 2000-01-07
; PRIOR APPLICATION NUMBER: PCT/AU98/00530
; PRIOR FILING DATE: 1998-07-09
; PRIOR APPLICATION NUMBER: AU P07765
; PRIOR FILING DATE: 1997-07-09
; PRIOR APPLICATION NUMBER: AU P09467
; PRIOR FILING DATE: 1997-09-11
; NUMBER OF SEQ ID NOS: 219
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Bovine
; OTHER INFORMATION: papillomavirus type 1 L2 open reading frame
; OTHER INFORMATION: (humanized)
; OTHER INFORMATION: Wild-type codons replaced with synonymous codons
; OTHER INFORMATION: used at relatively high frequency by human genes
US-09-479-645A-8

```

Query Match 17.0%; Score 408.5; DB 4; Length 469;  
 Best Local Similarity 29.3%; Pred. No. 3.8e-26;  
 Matches 152; Conservative 59; Mismatches 174; Indels 133; Gaps 24;

```

QY      8 RKRASAYDLKRTCKQAGCPDPVIRKVEGDTIAKILKFGGLATYLGGLGIGTWSGRV 63
Db      6 RKRASAYDLKRTCKQAGCPDPVIRKVEGDTIAKILKFGGLATYLGGLGIGTWSGRV 65
QY      64 GTGRTGYIPL--GGRSNTVVDVGP-----TRPPV-----VIEPVPTPST---VT 105
Db      66 AAGGSPRRYPLRTAGSTSSLASISRAVYAGTRPSIGAGIPLDTLETLGALRPGVEDTV 125
QY      106 LIESSVVTSGAPRPFTGTSGFDITSAGTTP-----AVLDIYP-SST 148
Db      126 LPEAPAIYTPDA----VPADSGLDALSTGDSSTETLITLLEPEGEDIAVLELQPLDRP 181
QY      149 SVSISTNTNPAFSDPSIIEVPQGEVSG--NVFVGPTSGTHGEELIPLQTFASSGCG 206
Db      182 TWQVNAVAHQSSAYHAPLQLO--SIAETSGLENIIEVGGSGLDGDTGGENIELTFGSG---- 236
QY      207 EEPISSTPLPYTRRVAGPRLY--SRAYQOVSVANPEFLRPSLIITYDPAPEPVDTTLT 264
Db      237 --PRISTPRSIASKSGILNMFSKRYTQVPTEDPEVF-----SSQTFANPLIEA----- 284
QY      265 FEPRSNVPSDFMDIIRLHRPALTSRGTVRFSRLGQRATMFTRSQTQIGARVHYHDIS 324
Db      285 -EP-----AVLKGPSGVRGLSQYKKPDLITRSGTEVGPQLHVRYSLS 326
QY      325 PIAPSPETI-----ELQPLVSATEDNGLEDIYADIDPA--MP----- 360
Db      327 TIEDEVALPYVDENTOGIAFVPL--HEEQAGFEIEILDSETHRLLPQNTSSTPGVS 384
QY      361 -----VPSR---PTTSSAVSTY-SPTISSASSYSNVYPLNLSMDVPYTYGDTILPPT 410
Db      365 GVRSLIPLTRESATRPCTGVVYTGSPDYTSAS-----PVTD---POSTSPSLVIDDT 433
QY      411 SVMPIVSPAPASTOYIGIHGTHYVLYLPIYFIPKRRK 448
Db      434 TTTPILII-----IDGHTVDLYSSNYTLHPSLLRKRRK 467

```

```

RESULT 12
US-08-484-791-2
; Sequence 2, Application US/08484791
; Patent No. 6380157
; GENERAL INFORMATION:
; APPLICANT: Jarrett, William F.H.
; APPLICANT: Campo, Maria S.
; APPLICANT: Smith, Kenneth T.
; TITLE OF INVENTION: Papillomavirus L2 Protein
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Flehr Hobach Test Albritton & Herbert LLP
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,791
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/170,185
; FILING DATE: 29-APR-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/GB92/01092
; FILING DATE: 17-JUN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9113809.9
; FILING DATE: 26-JUN-1991
; ATTORNEY/AGENT INFORMATION:

```





CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Coruzzi, Laura A.  
REGISTRATION NUMBER: 30,742  
REFERENCE/DOCKET NUMBER: 7682-037  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-8864  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 2  
SEQUENCE CHARACTERISTICS:  
LENGTH: 907 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-783-774-2

Query Match 6.3%; Score 151; DB 3; Length 907;  
Best Local Similarity 22.9%; Pred. No. 0.00042;  
Matches 115; Conservative 52; Mismatches 164; Indels 172; Gaps 25;

18 YKTCQSG--TCSDVNVKVEGTTLADKI-LQMSLIGFLGIGI-----GNGSGTG 67  
235 YESHVSQGLITSTSPVATPIPGTGYASRLPRPVSRGLNNSILYVYSGNGPKASG 294  
68 RIGYIPLGSRNTV-DVGPTRPVYIEVGPDPDSIVTLIEDS-----VWTS--GAP 118  
295 GDYCI-----OSNIVFSEIDIPASQDM-----PTNTDITYVGDNATVSPMVTSEDNASP 344  
119 RPTFTG-----TSG-----EDITSAGT--TTPAVLD 142  
345 NVTVIAFWAMPNNTETDFCKKWLITSGTPSGCENISGAFASNRTFDITVSGLTAPKTLI 404  
143 ITSPSTSVSIST-----TNFTNP-----AFSDPSITE--VPOTGEVSGNVFVGTPT 186  
405 ITRATNATTTTHKVIKSKAPESTTSPILNTGTFADPNTTGLPSTHYPTNLTAPAST 464  
187 SGHGYEIPLOTFPASSGTEBEPISSTPLPTVRRVAGPRLYSRAYQOVSVANPEFLTRPS 246  
465 GPTVSTADVTSPTPAGTTSASPVTPSPSPWDMGTE-----SKAPDMTSTSP--VYTPPT 517  
247 SLITYDNPAPEPVDITLTFPRSNVPSDMDIIRLHRPALTSRGRVRSRIGQRATMF 306  
518 PNATSPTRAV-----TTPPT--PNATSP-----TPAVTTPPNATSPPTLTK----- 555  
307 TRSGTQIGARVHYHDISPIAPSPXYIELQPLVSATEDNGLFDIYADDIDPAMPVPSRPT 366  
556 -----TSPTSAVTTPPNATSP-----PTLTKTS 578  
367 TSSAVST-----YSPTISSASSYSNVTPLTSSMDVPVYTGPD-----TLPP 409  
579 PTAIVTTPPNATSPPTLTKTSPTSAVTTTP-----TPNATGPTVGETSPQANATNHTLGG 632  
410 TSWPPIVS--PTAPASQYIGIH 430  
633 TSPTPVYTSQPKNATSAVTTGQH 655

RESULT 15  
US-09-328-599A-1  
Sequence 1, Application US/09328599A  
Patent No. 6432679  
GENERAL INFORMATION:  
APPLICANT: MOND, James J. and Tees, Andrew  
TITLE OF INVENTION: Enhancement of B Cell Activation by  
TITLE OF INVENTION: Co-ligation of Receptors for Antigen and Complement C3d  
TITLE OF INVENTION: Using EBV gp350/220 or EBV gp350/220 Peptide Adjuncts  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &  
ADDRESSEE: Dunnet, L.L.P.  
STREET: 1300 I Street, N.W., Suite 700

CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/328,599A  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fordis, Jean B.  
REGISTRATION NUMBER: 32,984  
REFERENCE/DOCKET NUMBER: 04995.6025-00000  
TELEPHONE: (202)408-4000  
TELEFAX: (202)408-4400  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 907 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-328-599A-1

Query Match 6.3%; Score 151; DB 4; Length 907;  
Best Local Similarity 22.9%; Pred. No. 0.00042;  
Matches 115; Conservative 52; Mismatches 164; Indels 172; Gaps 25;

18 YKTCQSG--TCSDVNVKVEGTTLADKI-LQMSLIGFLGIGI-----GNGSGTG 67  
235 YESHVSQGLITSTSPVATPIPGTGYASRLPRPVSRGLNNSILYVYSGNGPKASG 294  
68 RIGYIPLGSRNTV-DVGPTRPVYIEVGPDPDSIVTLIEDS-----VWTS--GAP 118  
295 GDYCI-----OSNIVFSEIDIPASQDM-----PTNTDITYVGDNATVSPMVTSEDNASP 344  
119 RPTFTG-----TSG-----EDITSAGT--TTPAVLD 142  
345 NVTVIAFWAMPNNTETDFCKKWLITSGTPSGCENISGAFASNRTFDITVSGLTAPKTLI 404  
143 ITSPSTSVSIST-----TNFTNP-----AFSDPSITE--VPOTGEVSGNVFVGTPT 186  
405 ITRATNATTTTHKVIKSKAPESTTSPILNTGTFADPNTTGLPSTHYPTNLTAPAST 464  
187 SGHGYEIPLOTFPASSGTEBEPISSTPLPTVRRVAGPRLYSRAYQOVSVANPEFLTRPS 246  
465 GPTVSTADVTSPTPAGTTSASPVTPSPSPWDMGTE-----SKAPDMTSTSP--VYTPPT 517  
247 SLITYDNPAPEPVDITLTFPRSNVPSDMDIIRLHRPALTSRGRVRSRIGQRATMF 306  
518 PNATSPTRAV-----TTPPT--PNATSP-----TPAVTTPPNATSPPTLTK----- 555  
307 TRSGTQIGARVHYHDISPIAPSPXYIELQPLVSATEDNGLFDIYADDIDPAMPVPSRPT 366  
556 -----TSPTSAVTTPPNATSP-----PTLTKTS 578  
367 TSSAVST-----YSPTISSASSYSNVTPLTSSMDVPVYTGPD-----TLPP 409  
579 PTAIVTTPPNATSPPTLTKTSPTSAVTTTP-----TPNATGPTVGETSPQANATNHTLGG 632  
410 TSWPPIVS--PTAPASQYIGIH 430  
633 TSPTPVYTSQPKNATSAVTTGQH 655

Search completed: October 17, 2003, 11:03:54  
Job time : 22.4783 secs

**THIS PAGE BLANK (USPTO)**